PB90-187204

Role of Grid-Based, Reactive Air Quality Modeling in Policy Analysis: Perspectives and Implications, as Drawn from a Case Study

Princeton Univ., NJ

Prepared for:

Environmental Protection Agency, Research Triangle Park, NC

Mar 90

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. 600/3-89/082	2.	3. RECIPIENT'S ACCESSION NO. PB90 187204/AS
4. TITLE AND SUBTITLE THE ROLE OF GRID-BASED, REACTIVE AIR QUALITY MODELING IN POLICY ANALYSIS: PERSPECTIVES AND IMPLICATIONS, AS DRAWN FROM A CASE STUDY		5. REPORT DATE March 1990 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Philip M. Roth[1], Charles E. Blanchard[2], and Steven D. Reynolds[3]		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS [1]21 Catskill Court, San Anselmo, CA 94960 [2]Center for Energy and Environmental Studies, Princeton University, Princeton, NJ 08544 [3]Systems Applications, Inc., San Rafael, CA 94903		10. PROGRAM ELEMENT NO. N104/C/06/06 6229/02 (FY-89) 11. CONTRACT/GRANT NO. Contract No. 68-01-6849
12. SPONSORING AGENCY NAME AND ADDRESS Atmospheric Research and Exposure Assessment Laboratory Office of Research and Development U.S. Environmental Protection Agency Research Triangle Park, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED Final 14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

EPA Project Report and Project Summary

16. ABSTRACT

A primary objective of this study is to improve the understanding of the role of performance quality in determining a model's acceptability and usefulness to the policy maker, and thus to aid in developing soundly-based expectations of the modeling process. The vehicle for pursuing this objective is examining the historical evolution of the Urban Airshed Model (UAM), a grid-based photochemical model whose basic formulation is similar to the Regional Acid Deposition Model (RADM), and its application to policy analysis in the South Coast Air Basin of California. A derivative objective is to draw implications from the findings to aid in appraising the merits of future pursuits, notably the mounting of comprehensive field programs to support the evaluation of regional acid deposition models.

In this report, we: (1) describe the UAM, (2) examine its predictive capability through scrutiny of historical performance statistics, (3) assess its degree of acceptance in the scientific and regulatory communities, based on information and viewpoints solicited through questionnaires and interviews, (4) examine the influence on the policy-making community of the UAM studies and, based on these "findings", (5) attempt to develop a perspective on the expectations of current regional acid deposition models.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT RELEASED TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

PB90-187204

EPA/600/3-89/082
March 1990

THE ROLE OF GRID-BASED, REACTIVE AIR QUALITY MODELING IN POLICY ANALYSIS:

PERSPECTIVES AND IMPLICATIONS, AS DRAWN FROM A CASE STUDY

by

Philip M. Roth
21 Catskill Court
San Anselmo, CA 94960

Charles L. Blanchard
Center for Energy and Environmental Studies
Princeton University
Princeton, NJ 08544

Steven D. Reynolds
Systems Applications, Incorporated
San Rafael, CA 94903

Contract No. 68-01-6849

Project Officer

Robin L. Dennis
Atmospheric Sciences Modeling Division
Atmospheric Research and Exposure Assessment Laboratory
Research Triangle Park, NC 27711

ATMOSPHERIC RESEARCH AND EXPOSURE ASSESSMENT LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

NOTICE

The information in this document has been funded wholly or in part by the United States Environmental Protection Agency under Contract No. 68-01-6849 to Program Resources, Incorporated. It has been subjected to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

During the 1980s the National Acid Precipitation Assessment Program supported development of the Regional Acid Deposition Model (RADM). While sound performance evaluation was to be a part of the development process, concern existed that policy makers may hold overly optimistic expectations of RADM's performance, and of the time to RADM's acceptability for unrestricted use. A primary objective of this study is to gain an improved understanding of the role of quality of performance in determining a model's acceptability and usefulness to the policy maker, and thus to aid in developing soundly-based expectations of the modeling process. The vehicle for pursuing this objective is examining the historical evolution of the Urban Airshed Model (UAM), a grid-based photochemical model that is similar in basic formulation to RADM, and its application to policy analysis in the South Coast Air Basin (SOCAB) of California. A derivative objective is to draw implications from the findings to aid in appraising the merits of future pursuits, notably the mounting of comprehensive field programs to support the evaluation of regional acid deposition models.

Analysis of time trends in performance statistics for UAM suggests that (1) the tendency toward underprediction of ozone concentrations decreased with time over fifteen years of testing, but the range in estimates of bias at a particular time were high and remained relatively unchanged with time, and (2) the average error in prediction decreased with time, but even in recent times displayed a significant range in variation.

Analysis of interviews of thirty individuals familiar with UAM use in policy analysis in the SOCAB suggests that (1) UAM's performance was "good enough for estimates to be considered" in the policy context, but not "good enough to provide a high degree of confidence ...", (2) increased acceptance of UAM was not related to improved performance (or the perception of improved performance) with time, (3) increased acceptance was related to improved understanding of the model's capabilities and limitations, (4) interviewees would accept model use in assessing the relative effects of alternative control scenarios, but not in quantitative evaluation of specific measures, and (5) interviewees favor increased emphasis on acquisition of better input data and more comprehensive data bases for supporting model evaluation.

In this report, we: (1) describe the UAM, (2) examine its predictive capability through scrutiny of historical performance statistics, (3) assess its degree of acceptance in the scientific and regulatory communities, based on information and viewpoints solicited through questionnaires and interviews, (4) examine the influence on the policy-making community of the UAM studies and, based on these "findings", (5) attempt to develop a perspective on the expectations to be placed on regional acid deposition models at the present time.

CONTENTS

Abstract iii
Figures vii
Tables vii
Acknowledgments viii

I. Introduction 1
  A. Motivation for the Study 1
  B. Perspective on the Study 2
  C. General Approach 3
  D. Structure of the Report 4

II. The Systems Applications, Inc. Urban Airshed Model 5
  A. Formulation of the Model 5
  B. History of Model Development and Application 7

III. The Predictive Performance of the Urban Airshed Model 11
  A. Range of Applications and Testing Conditions 11
  B. Trends in Quality of Performance 12
  C. Summary of Findings 26

IV. The UAM and Policy-Related Studies 29
  A. Policy Events 29
  B. Specific Questions of Concern 30
  C. Approach 31
  D. Survey Results 32
  E. Summary 41

V. The Influence of Model Results on Policy 43
  A. Factors Influencing Increased Acceptance of the UAM 43
  B. Assessing the Importance of Model "Competence" in Model Acceptance 45

CONTENTS (concluded)

VI. A Perspective on Evaluating RADM Performance: Lessons from the UAM Experience ... 46
  A. A "Model" of the Modeling Process ... 46
  B. Comparing the RADM and UAM Efforts and Experiences ... 47
  C. The Role of Performance Evaluation in Model Development ... 49
  D. Technical Acceptance of RADM ... 50
  E. The Acceptance of RADM in Policy Analysis ... 51

References ... 53

Appendices:

  A. Technical Features of the Systems Applications, Inc. Urban Airshed Model ... A-1
  B. Statistical Measures of Model Performance ... B-1
  C. Survey on Photochemical Air Quality Modeling ... C-1
  D. Summary of Interviews ... D-1

FIGURES

Number		Page
1	Time history of average bias in UAM O_3 results	17
2	Time history of average error in UAM O_3 results	18
3	Time history of average bias in UAM NO_2 results	19
4	Time history of average error in UAM NO_2 results	20
5	Bias in UAM predictions as a function of measured concentration for three Los Angeles studies--27 June 1974 simulation day	23
6	Absolute error in UAM predictions as a function of measured concentration for three Los Angeles studies--27 June 1974 simulation day	24
7	Distribution of distances to nearest bracketed value for O_3 for 27 June 1974	25
8	UAM O_3 predictions at Pasadena and Pomona for the 26-27 June 1974 episode	27
A-1	Schematic illustration of the grid used and treatment of atmospheric processes in the SAI Urban Airshed Model	A-2
A-2	Example of plots illustrating comparisons between computed and measured ozone concentrations (pphm) at air quality monitoring stations in the South Coast Air Basin for 8 and 9 August 1984.	A-6
A-3	Example of printed map of ground-level concentration predictions	A-7
A-4	Example of plotted contour map of ground-level concentration predictions	A-8

TABLES

Number		Page
1	Summary of Urban Airshed Model Performance Statistics	14
2	Number of Survey Respondents by Activity and Affiliation	32
3	Influence of UAM on Respondents' Assessments of Control Measures	33
4	Influence of UAM on Others' Assessments of Control Measures	34
5	Respondents' Assessments of UAM Performance Levels	35
6	Performance Measures Utilized by Respondents	37
7	Sensitivity Analyses Utilized by Interviewees	38
8	Factors That Would Increase Acceptance of the UAM	40

ACKNOWLEDGMENTS

The authors wish to acknowledge the support and counsel of Dr. Robin Dennis throughout the course of this study. They also express their appreciation to Messrs. Kenneth Schere and Joseph Tikvart for their review and helpful comments.

The authors also thank all those interviewed (see list at the end of Appendix C) for their time, cooperation, and valuable comments. Their participation was integral to this effort.

I. INTRODUCTION

A. MOTIVATION FOR THE STUDY

During the 1980s the National Acid Precipitation Assessment Program (NAPAP) has supported the development of the Regional Acid Deposition Model (RADM) from inception through planning for the evaluation of its predictive capabilities. Primary efforts have focused on:

- developing a comprehensive, soundly-formulated model;
- developing plans for its evaluation, including designing a field program for acquiring needed data and specifying the evaluative process itself; and
- subjectively estimating the "return on investment", that is, determining if the large expenditures entailed are likely to produce a satisfactory return.

While the technical and policy communities have spoken forcefully for the need for sound performance evaluation, the high costs of such an effort have given some pause. Of particular concern is the possibility that policy makers may hold overly optimistic expectations of performance. When funding of this magnitude is considered, the policy maker who may not be familiar with modeling may expect that "the evaluation project" -- conceived as a one-time event with monitoring to be conducted over two years -- will produce a "validated model"; in effect, a model that has a defined, acceptable level of uncertainty, and a model that can be used and depended upon in assessing policy alternatives.

A more realistic view is that modeling is a continuing process -- one of model development, design of an evaluative study (to stress the model, that is, to cause it to reveal its flaws if it is indeed inadequate), evaluation, diagnosis (of the causes of inadequate performance), and model improvement, followed by repetitions of the cycle. At any juncture a decision to apply a model is a decision to take advantage of a model's capabilities at its current state of development. Model application may serve one or more of several purposes: estimating impacts of specific control strategies; determining the expected changes in air quality under hypothesized emissions reductions scenarios; integrating and codifying current knowledge of emissions, transport, chemistry, and deposition; and identifying, through sensitivity studies, information and data acquisition needs.

If one envisions the sequence of steps comprising model development as a continuum in time, then a performance evaluation is a sampling of the process at a particular juncture. Evaluation does not mark an end of the process. The "sampling" of performance should not lead to the expectation of "adequate" prediction capability. Rather, the attainment of adequate performance is a goal to be pursued as deliberately and effectively as possible. But expectations should reflect the realities; the system under study is extremely complex and several evaluation cycles may be required.

This characterization of the modeling process may prove disappointing to: (1) policy makers, who need "good results" as soon as possible, and (2) nonscientists, who may view such an undertaking as being akin to an engineering project -- scope the effort, estimate costs properly, manage the project well, and the result will be a reliable, well-functioning product. Users of modeling results may expect that models are reliable, trustworthy surrogates of atmospheric processes. They may expect that a model's accuracy is known and is acceptably high. They may expect that the model's range of applicability is known, and that the model is used only within this range. Such expectations often are not fully met. Unless policy makers are well-informed about the nature of model development and the modeling process, they may hold unrealistic expectations in supporting a performance evaluation and in planning for a model's application. However, if performance evaluation is not undertaken, confidence in the model's predictive capability cannot be established, and the investment in model development will not provide a useful "return".

A potentially compelling means of developing a realistic portrayal of the model evaluation process is to examine the experiences of the past. In 1970, Systems Applications Inc. (SAI) was funded by the National Air Pollution Control Administration (NAPCA, a predecessor of the EPA) to develop a grid-based photochemical model of ozone formation on an urban scale -- the Urban Airshed Model (UAM). Now, eighteen years later, a subsequent version of that original model (in effect, a great-grandchild) is in use in a number of projects. Models, based on the same principles and deriving from the original version, have been developed for application on larger scales. In the intervening period of eighteen years, model improvement has been continuing and performance evaluation has been undertaken a number of times.

This history of model development, evaluation, and application is unique, particularly since it involves a complex, grid-based "chemical" model similar in basic formulation to RADM (although UAM applies only for gas phase species and over a much more restricted spatial scale, thereby giving less emphasis to deposition processes and no attention to cloud and scavenging processes). Moreover, model development and evaluation has proceeded in time as a scientific undertaking under the scrutiny of the regulatory community. Thus, the examination of UAM's evolutionary history provides a unique opportunity to learn. The limited assessment reported here, then, offers insight gained from one case study that we hope will prove useful to the policy maker in developing soundly based expectations of the models and the modeling process.

B. PERSPECTIVE ON THE STUDY

A goal of this study is to develop a realistic perspective concerning expectations of performance evaluation -- and to do so, in part, through examining the history of the development and multiple evaluations of a somewhat similar model. This report focuses on the process of evaluation, modeling achievements, the time and the number of iterations apparently required for successful evaluation, the limitations of models and their predictive performance, and the circumstances that influence model acceptance and use.

As noted, performance evaluation may be viewed as an episode in the continuing evolution of a model and its development. However, the political climate today is placing great pressure on the developers of acid deposition models to produce models displaying "satisfactory performance" under intense scrutiny, in a tight time frame, and for predetermined costs. Because of the nature of model development, it may not be possible to meet the various expectations -- particularly those held by "funders" and decision makers -- Congress, the EPA, the Office of Management and Budget (OMB), and other agencies.

Yet model development and application is an essential component in the overall process of assessing national emissions control strategies. The cost of acquiring the data base needed for the evaluation of a model of regional or subcontinental scope is very high. Expectations tend to be correspondingly high for such a project, as compared with less costly undertakings. With an investment of tens of millions of dollars, policy makers want to be reasonably certain that a useful return will be realized.

C. GENERAL APPROACH

The primary objective of the study is to gain an improved understanding of the importance of quality of performance in determining a model's acceptability and usefulness to the policy maker. A vehicle for pursuing this objective is to examine the influence of modeling in policy analysis through a study of the historical evolution of the UAM, as applied in the South Coast Air Basin (SOCAB) of California over a period of years. By virtue of the several evaluations and many applications carried out over the past decade or so, this use of one model provides a uniquely rich history for examination.

A derivative objective is to draw implications from the findings that may be of value in appraising the merits of future pursuits, notably the expenditure of large sums of money on field programs to support the evaluation of regional acid deposition models.

In this study we surveyed both the performance evaluation history of the UAM and the degree of acceptance achieved by the model in the policy making and scientific communities. We first examined and interpreted performance statistics for the model over the history of its development. We then interviewed three categories of interested and informed parties -- members of the contemporaneous scientific community whose pursuits led them to become familiar with the UAM and its development history, industry representatives, and regulatory agency staff and policy makers who have had to assess the merits of using the results of the UAM. Based on the findings of these efforts, we developed a "subjective history" of the perception of model performance and of model acceptance.

A critical factor in the development of grid-based urban airshed models and acid deposition models is the level of commitment accorded each in providing designed data bases to support model evaluation and application. During the fifteen year development and application period of the UAM (through 1985), virtually all evaluation efforts were dependent on the use of routine monitoring data, sometimes supplemented by additional measurements. An important exception is the Regional Air Pollution Study (RAPS), during which a comprehensive data base was collected in St. Louis to support model development and application (Schiermeier, 1978). This relatively limited support of model development may be compared with the monitoring program being implemented for the evaluation of RADM.

[In September 1985, a comprehensive, designed data base was collected for the three contiguous California counties of Santa Barbara, Ventura, and San Luis Obispo under the South Central Coast Cooperative Aerometric Monitoring Program (SCCCAMP) specifically for use in grid model performance evaluation. Evaluation studies are now in progress. During the summer of 1987, a second data collection program in support of model development and application, the South Coast Air Quality Study (SCAQS), was carried out in the Los Angeles air basin. A third, more limited data base was collected for Kern County during the summer of 1984. Even though the purpose of the Kern study was to acquire aerometric data aloft to aid in designing a future, more extensive program, it nevertheless proved most useful for model performance evaluation. These three studies provided the primary "evaluation quality" data bases for use in testing grid-based models that have been collected to date in California. (Of these studies only the results of the Kern effort were available for inclusion in this analysis.) While monitoring studies were carried out earlier in St. Louis (RAPS), Tulsa, and Philadelphia to support performance evaluation, California data are needed to evaluate the model under the terrain, meteorological, air quality, and emissions conditions that prevail in the metropolitan regions of the state.]

D. STRUCTURE OF THE REPORT

In Chapter II of this report we describe the UAM, the model whose performance is the subject of scrutiny in this study. Chapter III is devoted to an examination of the predictive performance of the UAM over the period of the mid-1970s through the mid-1980s. In assessing performance characteristics statistical measures used in the past are presented and analyzed. In Chapter IV, acceptability of the UAM in the scientific and regulatory communities is assessed, based on information and viewpoints solicited from "the communities of interest" through questionnaires and interviews. The period of exposure to the model and its application of those individuals contacted coincides with the period of performance evaluation.

In Chapters V and VI, we examine the influence on the policy making community of the UAM studies and, based on these "findings", attempt to develop a perspective on the expectations to be placed on the evaluation and early use of regional acid deposition models.

II. THE SYSTEMS APPLICATIONS, INC. URBAN AIRSHED MODEL

Efforts to develop the Systems Applications, Inc. (SAI) Urban Airshed Model (UAM) were initiated in 1970 under contract to the NAPCA, which in 1971 was incorporated into the EPA. At that time, the EPA was faced with the need to identify the emission controls required in designated urban areas to meet the National Ambient Air Quality Standard for photochemical oxidants promulgated in 1967. Furthermore, air quality models at that time were based on Gaussian formulations which did not provide a means for treating nonlinear chemical reaction processes such as those leading to oxidant formation. Thus, in an attempt to provide suitable photochemical modeling capability, the EPA sponsored three independent developmental studies. Investigators at General Research Corporation (Eschenroeder, Martinez, and Nordsieck, 1973) and Pacific Environmental Services (PES, 1973) formulated trajectory models and applied them to the simulation of six smoggy days that occurred in Los Angeles in 1969. Systems Applications, Inc. was selected to develop a grid-based approach (i.e., the UAM) and apply it to the same smog episodes. The EPA evaluated the results of the three developmental studies in 1973 and decided to limit further support to the continued development of the grid-based UAM.

Photochemical models simulate ozone formation. Ambient air quality monitoring methods and standards in the late 1960s and early 1970s provided or required information concerning photochemical oxidants. Although ozone is the primary component of oxidants in most areas, photochemical models were not able to provide concentration predictions (ozone) consistent with the available measurements and standards (oxidants). This situation was rectified in the late 1970s when the EPA's reference measurement method and the National Ambient Air Quality Standards were revised in terms of ozone, thus placing measurements, standards, and photochemical models on the same basis.

A. FORMULATION OF THE MODEL

The formulation of the UAM is based on the time-dependent, three dimensional equations describing the conservation of mass of a chemically reactive pollutant in a turbulent atmosphere. In order to solve the resulting system of coupled nonlinear partial differential equations, it is necessary to utilize numerical integration techniques, which further require that the modeling region be partitioned into an array of three dimensional grid cells. These cells have fixed horizontal dimensions of 2 km to 10 km and time-varying vertical dimensions of 50 meters to several hundred meters. The overall modeling region may extend from 50 km to 200 km in the horizontal and from less than 1 km to 3 km in the vertical. Ideally, the geographical extent of this region should contain all areas with sources of reactive hydrocarbon and NO_x precursor emissions that contribute to ozone formation and the downwind areas where significant ozone levels are likely to occur.

The UAM includes treatments of the important physical and chemical processes that contribute to ozone formation in and immediately downwind of urban areas. In particular, the model provides a representation of the following phenomena:

- Precursor emissions. The spatial and temporal characteristics of reactive hydrocarbon and NO_x emissions sources must be supplied as inputs to the model. Emissions from area and small point sources are treated as volume sources emitted into the appropriate ground-level grid cells as a function of time. Emissions from large point sources are injected into upper-level grid cells based on the calculated plume rise. Hydrocarbon emissions must be apportioned into groups (e.g., paraffins, olefins, aromatics, etc.) according to the speciation requirements of the chemical kinetic mechanism embedded in the model.

- Pollutant transport. Once the ozone precursors are emitted into the atmosphere, they are transported by the wind. When ozone is formed, it is also subject to transport by the wind. The model requires the preparation of three dimensional, time varying estimates of the hourly-averaged wind speed and direction. These values must be specified for each grid cell.

- Turbulent diffusion. Ozone and its precursors are also subject to turbulence-related dispersion processes that take place on a subgrid scale. These turbulent diffusion effects are represented using the so-called gradient transport hypothesis, where the pollutant flux is assumed to be proportional to the spatial gradient in the concentration field. The turbulent diffusivities employed in the model are dependent on atmospheric stability and other meteorological variables.

- Chemical reactions. Ozone concentrations above ambient background levels are the result of chemical transformations involving reactive hydrocarbons and NO_x. A chemical kinetic mechanism representing the important reactions that occur in the atmosphere is employed to estimate the net rate of formation of each pollutant simulated by the model. The mechanism includes provisions for treating differences in reactivity of various hydrocarbons as well as the spatial and temporal variations in ultraviolet light intensity, which influences the rate at which photolysis reactions take place.

- Removal processes. Pollutants are removed from the atmosphere via interactions with surfaces at the ground. The rate of removal for each pollutant is determined by two factors: 1) the estimated rate of transport to the surface (influenced largely by the wind speed and stability near the surface), and 2) the rate of uptake at the surface (influenced by the land-use characteristics in each ground-level grid cell and the rate at which various types of vegetation and building materials adsorb pollutants).

- Initial and boundary concentrations. To solve the governing equations of the model, the concentration of each pollutant in each grid cell must be specified at the start of the simulation. Furthermore, the transport of ozone and precursors into the modeling region from upwind areas is represented in the boundary concentration inputs, which may vary both spatially and temporally.

An important attribute of the UAM's formulation is the framework it provides within which to represent the current understanding of the physical and chemical processes that contribute to ozone formation in and downwind of urban areas. Thus, as improved representations of atmospheric phenomena are developed, existing algorithms or process descriptions can be replaced or added by modifying the appropriate portions of the computer code.

B. HISTORY OF MODEL DEVELOPMENT AND APPLICATIONS

Initial UAM developmental efforts and the application to Los Angeles to simulate six single day ozone events in 1969 are fully described in the two sets of reports by Roth et al. (1971) and Reynolds et al. (1973a). These reports are summarized in Reynolds et al. (1973b, 1974) and Roth et al. (1974). Evaluation of the model's performance and its components at that time showed that improved treatments of some physical and chemical processes were necessary. Sponsored by the EPA, another series of research efforts was carried out between 1973 and 1977 to improve the model. Notable modifications to the model may be summarized as follows:

- Replacement of the 15-step Hecht and Seinfeld (1972) kinetic mechanism with the Hecht, Seinfeld, and Dodge (1974) mechanism, which was subsequently replaced by the initial version of the Carbon-Bond Mechanism (CBM-I) developed by Whitten and Hogo (1977). In addition, the effects of temperature and aerosol scattering of ultraviolet light on reaction rate constants, a computationally-efficient representation of nighttime ozone chemistry, and a simple mechanism describing homogeneous SO_2 oxidation were also included.

- Incorporation of an improved numerical integration procedure based on the Sharp and Smooth Transport Algorithm (SHASTA).

- Development and inclusion of improved representations of vertical diffusivity, plume rise, and surface removal processes.

- Development of a module to treat the influence of subgrid scale NO_x emissions from point and line sources on the grid-volume-average ozone concentrations generated by the model.

- Modification of the computer codes to enable the input of a fully three-dimensional wind field.

These developmental activities are documented in a seven-volume series of reports by Jerskey and Seinfeld (1977), Jerskey et al. (1976), Killus et al. (1977), Lamb (1976), Lamb et al. (1977), Liu et al. (1976a), and Reynolds et al. (1976), and in the papers by Lamb, Chen and Seinfeld (1975), Liu and Seinfeld (1975), Liu, Whitney and Roth (1976b), and Reynolds (1977).

Improvements to the software and documentation as well as application of the model to St. Louis using data collected as part of the Regional Air Pollution Study (RAPS) were supported by the EPA during the period from 1977 through 1979. In restructuring the computer codes, provisions were included to enable the simulation of multiple day episodes and the inclusion of large numbers of grid cells through the concept of grid segmentation (i.e., partitioning the modeling region into subregions). Two additional developmental studies of note include one sponsored by the EPA to provide an improved representation of reactive plume dispersion at the subgrid scale, which lead to the development of the Plume-Airshed Reactive Interactive System (PARIS), as described by Seigneur et al. (1983), and ongoing work supported by the South Coast Air Quality Management District aimed at the implementation of improved aerosol predictive capabilities.

During the early 1970s, there was little support for application of the UAM to urban areas other than Los Angeles or to evaluate possible emission control strategies. The first applications were quite limited in scope. One involved the simulation of four motor vehicle control strategies for Los Angeles (Reynolds et al., 1973a), the other an evaluation of two future year emissions scenarios associated with the proposed State Implementation Plan of June 1973 (Reynolds and Seinfeld, 1975). Both studies were minor adjuncts to ongoing developmental efforts.

The first serious interest in applying the UAM was expressed by the California Department of Transportation (CALTRANS) and the California Air Resources Board (CARB), who jointly sponsored a study in 1974 to expand the original SOCAB modeling region of Reynolds et al. (1973a) from 50 x 50 miles to 80 x 160 miles and to establish a set of model inputs for 26 and 27 June 1974 (the highest ozone concentrations observed in 1974 occurred on the latter day). Serious budgetary problems encountered by CALTRANS during this study ultimately limited its ability to continue supporting the activity. Nevertheless, the inputs developed at that time (Ross, 1975; Tesche and Whitney, 1975) were to be the subject of significant review and scrutiny and served as a point of departure for several UAM applications to the SOCAB over the next ten years.

At about the same time, the State of Colorado Division of Highways funded a study to apply the 1973 version of the UAM to Denver in order to show that various transportation plans and programs were consistent with the SIP, as required by the 1970 Federal-Aid Highway Act. In a paper summarizing the early work of the Division of Highways, Donnelly (1978) noted that the model was used to prepare a 109(j) air quality assessment statement (Colorado Division of Highways, 1976a), an air quality maintenance plan statement (Colorado Division of Highways, 1976b), and an environmental impact statement for the proposed I-470 beltline interstate project (Colorado Department of Health, 1977). Subsequently, the EPA Region VIII sponsored a study (Anderson et al., 1977) to apply the UAM to three smoggy days in 1975 and 1976 and to evaluate the effects on Denver's air quality of urban growth and development that would potentially result from the implementation of proposed wastewater treatment facilities.

In July 1974, Systems Applications was awarded a contract by the U.S. Department of Transportation, Federal Highway Administration (FHWA) to develop and refine a set of photochemical models to be used by transportation planners and highway engineers to examine the effects of transportation control strategies, to evaluate the performance of the models, and to assess the potential effectiveness of highway-related control measures as a means of reducing photochemical smog formation. As documented in the report by Reynolds et al. (1979), single and multiple day simulations were performed for the 26 and 27 June 1974 episode period in the SOCAB, the three 1975 and 1976 smoggy days in Denver, and two simulation days in Sacramento. Several sensitivity runs were carried out for a "hypothetical" city to examine the possible influence of alternative highway location, design, and operation strategies on ozone formation.

During the course of the FHWA investigation, the Southern California Edison Company (SCE) sponsored a study to examine the applicability of the UAM to simulate the air quality impacts of selected SCE power plants in the SOCAB. Of particular interest was the use of the model to examine issues associated with the implementation of NO_x emission controls on power plants in the SOCAB. Because of the complexity of the 26-27 June 1974 meteorological conditions, Tesche and Burton (1978) argued that there was a need to select another modeling period to adequately assess power plant impacts. They selected the 4 August 1975 simulation day. In addition to evaluating model performance, Tesche and Burton (1978) and Tesche, Burton, and Mirabella (1979) reported the results of two simulations, the first involving a 90 percent reduction in NO_x and SO_2 emissions from power plants, and the second, a 50 percent reduction in hydrocarbon and NO_x emissions from motor vehicles.

The second phase of the SCE work focused on the simulating a multiple day episode, largely to avoid problems connected with specifying initial concentration inputs for single day runs when performing future year control strategy simulations. Based on the findings of concurrent studies, the 26-27 June 1974 period was selected for use. During this study, the second version of the Carbon-Bond Mechanism (CBM-II) was included in the UAM. In addition, further consideration was given to specifying model inputs, in particular, boundary concentrations and the removal of areas of excessive convergence and divergence in the three dimensional wind fields, as discussed by Killus et al. (1981). An important product of the study was an assessment of the impacts of NO_x controls, as specified by Rule 1135.1 of the South Coast Air Quality Management District and Rule 59.1 of the Ventura County Air Pollution Control District. Using projected 1987 emissions estimates, Killus et al. (1981) used the UAM to examine the effects on ozone formation of 90 and 50 percent reductions in power plant emissions in the SOCAB.

Four other important UAM applications to the SOCAB were conducted in parallel with the SCE work described above. For the Southern California Association of Governments (SCAG), simulations were performed using the meteorological conditions of 4 August 1975 to study the influence of six alternative population, housing, employment, and land use forecasts for 1990 on predicted ozone concentrations (Reynolds et al., 1978). The EPA Office of Air Quality Planning and Standards sponsored a large sensitivity study to provide guidance concerning the types and amounts of data required to support the application of the UAM to an urban area. Simulations were performed using the meteorological conditions of 26 June 1974 and 4 August 1975; the results of this effort are reported by Tesche et al. (1981a). The third study was carried out for the National Commission on Air Quality (NCAQ) and involved an assessment of the ambient air quality impacts of future year emissions resulting from implementing alternative control technologies and population and industrial growth scenarios in the SOCAB. Multiple day simulations using the meteorological conditions of 26-27 June 1974 were employed to study four 1987 and one 1992 projected emissions scenarios. The results of the NCAQ study are documented in the report by Souten, Tesche and Oliver (1981). Finally, the California Air Resources Board commissioned several model sensitivity runs to examine the effects of several reactive hydrocarbon and NO_x emission reductions on predicted ozone and NO_2 concentrations (Tesche et al., 1981b).

The most recent major application of the UAM to the SOCAB was sponsored by the Western Oil and Gas Association to examine the possible effects of NO_x controls on one-hour average ozone and NO_2 concentrations. A review of NO_2 results from previous simulations of 26-27 June 1974 indicated a need to improve the model's ability to predict peak NO_2 concentrations that are observed during the mid-morning hours. This was accomplished based on a review and subsequent revision of mixing height inputs during the morning hours. In order to provide information pertaining to the effectiveness of control strategies during NO_2 episode conditions, which typically occur in the late fall and winter months, a complete set of model inputs were prepared for the NO_2 episode of 7-8 November 1978. Model input preparation and performance results are summarized in the reports by Tesche et al. (1983a, 1983b) and Oliver, Hogo and Saxena (1983). Several emissions sensitivity runs were made to examine the relative effectiveness of reactive hydrocarbon and NO_x controls in reducing ozone and NO_2 concentrations for both summer and fall episode conditions. This study was intended to parallel and supplement efforts by the public agencies in preparing the 1982 revision to the Air Quality Management Plan. The sensitivity results are documented in the reports by Seigneur et al. (1983b) and Roth et al. (1983) and in the paper by Roth et al. (1984).

Current interest in UAM applications in the SOCAB is being devoted to the incorporation of improved aerosol prediction capabilities, motivated largely by the recent adoption of state and federal PM-10 standards. These studies are being supported by SCE and the South Coast Air Quality Management District.

While this discussion of UAM developmental and applications experience has focused primarily on the SOCAB, a significant body of experience has also been acquired in applications to other areas in the U.S. and Europe. We have cited earlier applications to Denver, Sacramento, and St. Louis. The EPA has sponsored applications to Tulsa (Reynolds et al., 1982) and Philadelphia (Haney and Braverman, 1985). The Joint Interagency Modeling Study (JIMS) sponsored by several local, state, and federal regulatory agencies, has involved a large effort to model the South Central Coast Air Basin in California (including portions of San Luis Obispo, Santa Barbara, and Ventura Counties). This work is documented in the report by Haney et al. (1986). Extension of the JIMS applications experience to examine issues concerning petroleum developments offshore of San Luis Obispo County have been carried out by Yocke et al. (1985). In the last two years, considerable work has been performed by Systems Applications and CARB to apply the UAM to assess emission control issues in Kern County (CARB, 1985; Whitten et al., 1986). In Europe, the UAM has been applied to a large portion of The Netherlands (Builtjes and Reynolds, 1984), the Rhine-Ruhr area of West Germany (Stern and Scherer, 1984), and Turin, Italy.

III. THE PREDICTIVE PERFORMANCE OF THE URBAN AIRSHED MODEL

Prior to applying the UAM to assess the effects of proposed emission controls or other technical or regulatory issues, it is necessary to evaluate its ability to adequately represent the physical and chemical phenomena that take place in the modeling region. This performance evaluation is carried out by using the model to simulate one or, preferably, more historical episodes for the pollutant(s) of interest. Available emissions, meteorological, air quality, and other pertinent data are analyzed and then used to prepare model inputs that recreate as accurately as possible the conditions that actually occurred during each episode. In simulating photochemical air quality, calculated ozone, NO_2, NO, reactive hydrocarbon, and CO concentrations are compared with available measured values to assess the adequacy of the model's performance.

A. RANGE OF APPLICATIONS AND TESTING CONDITIONS

As indicated in the previous section, the UAM has been applied to several urban areas in the U.S. and Europe, and most of these studies have included some form of performance evaluation. Thus, there is a growing body of information concerning the accuracy of the model's predictions. Of particular interest to this investigation is the evolution of model performance over time, especially in Los Angeles applications since the model has been used most extensively in this area. In addition, the model's ability to simulate atmospheric processes in other areas provides further insight into its range of applicability.

Many applications of the model have involved single-day simulations. That is, the modeling period typically starts during the early morning hours (prior to 6 a.m.) and ends at some time later that day after the peak ozone concentrations were observed. A particular limitation of such simulations is that initial concentration inputs for ozone and its precursors may have a significant influence on the predicted ozone levels for several hours after the start of the run. The uncertainties in these inputs can be quite large (due to the absence of appropriate measurements), thus introducing significant uncertainties in the ozone results over the simulation period. Even more important, it is difficult to accurately specify initial concentration inputs for simulations of future-year periods, where the magnitude and spatial pattern of emissions may be significantly different from present baseline conditions. To avoid these limitations, model runs in more recent studies have been designed to cover multiple-day periods so that calculated ozone levels during the times of particular interest are only minimally affected by the initial concentration inputs.

Early model performance results for Los Angeles were for single-day simulations. Since 1979, however, most reported results have been for multiple-day runs. In other study areas, multiple-day simulations have been carried out for the South Central Coast Air Basin, San Luis Obispo, and Kern

County. All other applications, including those to Sacramento, Denver, Tulsa, and Philadelphia involve single-day runs. Thus, care must be exercised in characterizing and interpreting historical trends in model performance, taking into account differences that may be due to the length of the simulation period.

To evaluate model performance, calculated ozone, NO, NO_2, hydrocarbon, and CO concentrations are examined in light of the available measurements for each respective constituent. Unfortunately, there are currently no widely-accepted, standardized measures for representing model performance, although several studies have been carried out to develop and recommend such measures (e.g., Hayes, 1979; Bencala and Seinfeld, 1979; Fox, 1981; Dennis et al., 1983). Particular interest has focussed on the model's ability to predict relatively high measured one-hour average ozone concentrations since ambient air quality standards are compared with these values. Other attributes of interest are measures of bias and (absolute) error. Correlations of calculated and measured values as well as measures to assess the degree of agreement between the frequency distributions of the measured and predicted values have also been reported.

In early model applications (e.g., Reynolds et al., 1973a), it was necessary to compare predicted ozone concentrations with available observations of oxidant. Although ozone is the primary constituent of oxidant in most urban areas, it was not until the late 1970s that the National Ambient Air Quality Standards and the EPA's reference measurement methods were revised in terms of ozone, thus placing the model's predictions, ambient measurements, and the standard on a common basis.

The model's ability to accurately simulate CO (essentially an inert pollutant) has been used as a test of the emissions, meteorological and dispersion components of the model. However, monitoring stations in some areas are located near major roadways and intersections and thus do not provide an accurate observation of the volume-averaged CO concentrations calculated by the model. Similarly, an assessment of model performance for NO can be hindered by the proximity of monitoring stations to major roadways and stationary sources. Since model performance results for CO and NO have been reported in only a few studies (e.g., Reynolds et al., 1973a, 1979), they will not be considered further here. The dearth of available ambient hydrocarbon measurements makes it difficult to assess the model's performance for these important ozone and NO_2 precursors. Thus, our assessment of historical trends in model performance will be limited to consideration of ozone and NO_2.

B. TRENDS IN QUALITY OF PERFORMANCE

To characterize the level of performance achieved in Los Angeles and other areas, we have reviewed several reports published over the past 15 years documenting the most significant UAM applications. A difficulty in developing a clear statement of the historical trends arises as a result of inconsistencies in the procedures employed by different investigators to analyze model performance. More generally, comparison between any pair of studies included in this review of UAM performance may display one or more of the following inconsistencies:

- differing geographical areas (reflecting differing terrain and meteorology);

- differing quality, amounts, and ranges in coverage of meteorological and air quality data;

- differing lengths of simulation periods, i.e., single day episode versus second day of a two day episode versus first day of a two day episode (where performance in this last circumstance would generally be ignored today);

- differing lists of pollutants studied in performance evaluation;

- differing modes of construction and quality of emissions inputs; and

- differing calculational and reporting procedures.

Yet, in this report we attempt to analyze and compare performance measures in the hope that the results will be revealing and useful.

Initial attempts by Reynolds et al. (1973a, 1974) to characterize performance relied largely on the production of time series plots and spatial maps of calculated and measured concentrations and a qualitative discussion of the findings. Liu et al. (1976a) were the first to carry out a more systematic statistical analysis of these results, although their findings cannot be directly compared with those developed in more recent studies. The performance measures developed by Hayes (1979) have been adopted in several UAM studies and provide the most consistent basis for establishing historical trends. Table 1 provides a summary of the overall bias and error in simulation results reported for Los Angeles and several other urban areas. The bias for a particular pair of predicted and observed values is expressed as the difference in values (predicted minus observed) divided by the observed value. This fractional value is calculated for all pairs for which the observed value exceeds a specified threshold value (generally from 0.02 to 0.05 ppm, depending on the study).

Overall bias, the average of all fractional differences for a particular simulation day, is reported in Table 1. As an example, the bias in ozone results for the 9 September 1969 simulation day is -0.348. That is, the model tends to underpredict the measured values on this day by 34.8 percent. Overall error is calculated in a similar manner, except that the absolute value of each fraction is employed in generating an average instead of the signed value. Thus, the (absolute) error is greater than or equal to the bias because positive and negative values offset each other in calculating the average value for the bias. A further discussion of these performance measures is given in Appendix B.

The results cited in Table 1 for the six simulation days in 1969 were derived for use in the present investigation by digitizing time series plots of predicted and measured ozone and NO_2 concentrations as reported by Reynolds et al. (1973a). Also note that the initial 4 August 1975 simulation for Los Angeles was carried out by Tesche and Burton (1978). The simulation results were subsequently analyzed by Reynolds et al. (1979), who reported the bias and error figures presented in Table 1.

TABLE 1. SUMMARY OF URBAN AIRSHED MODEL PERFORMANCE STATISTICS

Urban area/ Simulation day	Ozone bias	Ozone error	NO_2 bias	NO_2 error	Reference
Los Angeles					
9 Sep 69	-0.348	0.645	-0.029	0.270	Reynolds et al. (1973a)
29 Sep 69	-0.177	0.366	0.123	0.330	
30 Sep 69	-0.451	0.611	0.114	0.379	
29 Oct 69	-0.006	0.557	0.183	0.284	
30 Oct 69	-0.364	0.639	0.177	0.356	
4 Nov 69	-0.370	0.536	0.174	0.349	
4 Aug 75	-0.297	0.569	-0.105	0.504	Tesche and Burton (1978)
26 Jun 74	-0.064	0.721	-0.240	0.472	Reynolds et al. (1979)
27 Jun 74	-0.430	0.650	-0.209	0.639	
26 Jun 74	-0.060	0.344			Tesche et al. (1981)
4 Aug 75	-0.366	0.404			
26 Jun 74	0.242	0.436			Killus et al. (1981)
27 Jun 74	0.103	0.373			
26 Jun 74	-0.050	0.370	-0.287	0.361	Tesche et al. (1983a)
27 Jun 74	0.100	0.330	-0.390	0.456	
7 Nov 78	-0.099	0.429	-0.109	0.373	Tesche et al. (1983b)
8 Nov 78	-0.212	0.547	-0.077	0.365	
8 Aug 84	0.034	0.216	-0.198	0.443	Hogo and Yocke (1987)
9 Aug 84	0.102	0.283	0.099	0.499	
South Central Coast					
11 Sep 80	-0.106	0.383			Haney et al. (1986)
12 Sep 80	-0.322	0.378			
25 Sep 80	-0.106	0.351			
26 Sep 80	-0.081	0.348			
28 Sep 80	-0.149	0.261			
29 Sep 80	-0.152	0.278			

(continued)

TABLE 1 (concluded).

Urban area/ Simulation day	Ozone bias	Ozone error	NO_2 bias	NO_2 error	Reference
San Luis Obisbo					
12 Sep 83	-0.017	0.224			Yocke et al. (1986)
27 Sep 84	0.074	0.306			
Kern County					
7 Aug 84	-0.028	0.296			CARB (1985)
8 Aug 84	-0.186	0.311			
9 Aug 84	-0.116	0.267			
10 Aug 84	0.000	0.237			
Sacramento					
28 Jun 76	-0.008	0.223	-0.395	0.448	Reynolds et al. (1979)
24 Aug 76	-0.224	0.287	-0.627	0.627	
Denver					
29 Jul 75	-0.086	0.502	-0.142	0.277	Reynolds et al. (1979)
28 Jul 76	-0.150	0.571	-0.500	0.500	
Tulsa					
29 Jul 77	-0.070	0.210	0.180	0.390	Reynolds et al. (1982)
3 Aug 77	-0.040	0.310	-0.100	0.420	
Philadelphia					
13 Jul 79	0.280	0.340			Haney and Braverman (1985)
19 Jul 79	0.080	0.310			

In preparing our tabulations of model performance, we wish to acknowledge the work of Haney et al. (1986), which provides a summary of results for several UAM applications.

To assist in the characterization of historical trends in model performance, we have plotted the bias and error values from Table 1 as a function of the year they were published. Figures 1 through 4 summarize these results for ozone and NO_2. In addition, we have used different symbols to help distinguish results for Los Angeles, the California South Central Coast Air Basin (i.e., the Santa Barbara, Ventura, and San Luis Obispo County areas), and other urban areas. Note that applications to the South Central Coast Air Basin present unique modeling problems associated with the need to represent complex coastal meteorological phenomena not present in the applications to the other inland urban areas.

A review of the bias results for ozone presented in Table 1 and Figure 1 yields the following observations:

- For most studies considered here, the model tends to underpredict.
- The first version of the model tended to underpredict the observed ozone values by up to 45 percent.
- The underprediction tendency for ozone has decreased over the years; in fact, the most recent results for Los Angeles indicate an overprediction tendency. Most results published since 1985 indicate a bias range of from +10 to -15 percent.
- Multiple-day simulation results have also improved over the years to the point that the bias in such results is comparable to that of single-day results.

The historical trend in the overall (absolute) error for ozone is shown in Figure 2. These results may be summarized as follows:

- Discrepancies between predicted and measured ozone values in early applications of the UAM to Los Angeles and Denver were about 50 to 70 percent.
- Discrepancies between predicted and measured ozone values have generally decreased over time; results since 1985 indicate a level of agreement between predicted and measured values to within about 20 to 40 percent of the observations.
- There is a clear indication that continued study of a particular episode yields improvements in performance. For example, the efforts of Killus et al. (1981) have led to significant improvements in model performance for the 26-27 June 1974 ozone episode over that achieved in the initial UAM study of this episode reported by Reynolds et al. (1979). Further efforts by Tesche et al. (1983a) yielded only modest

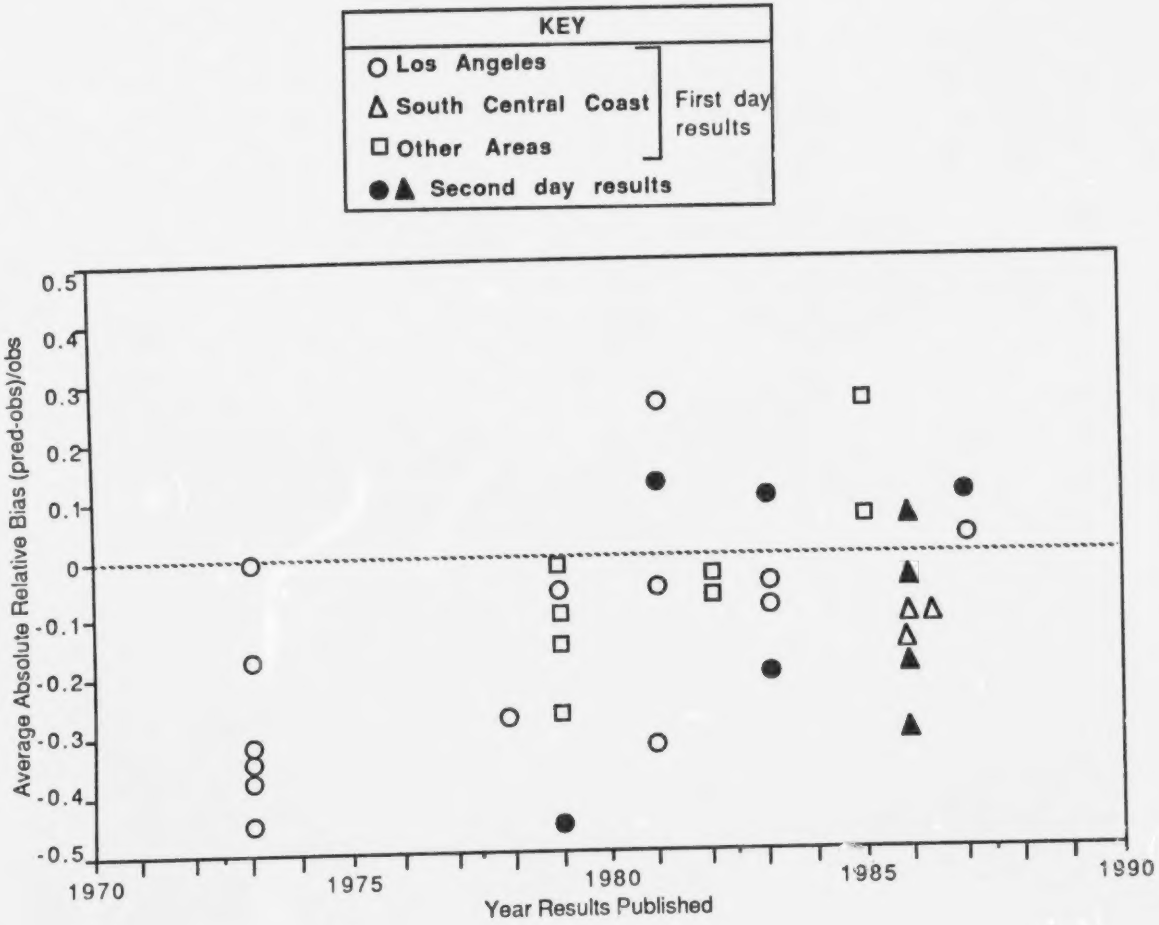


Figure 1. Time history of average bias in UAM O_3 results.

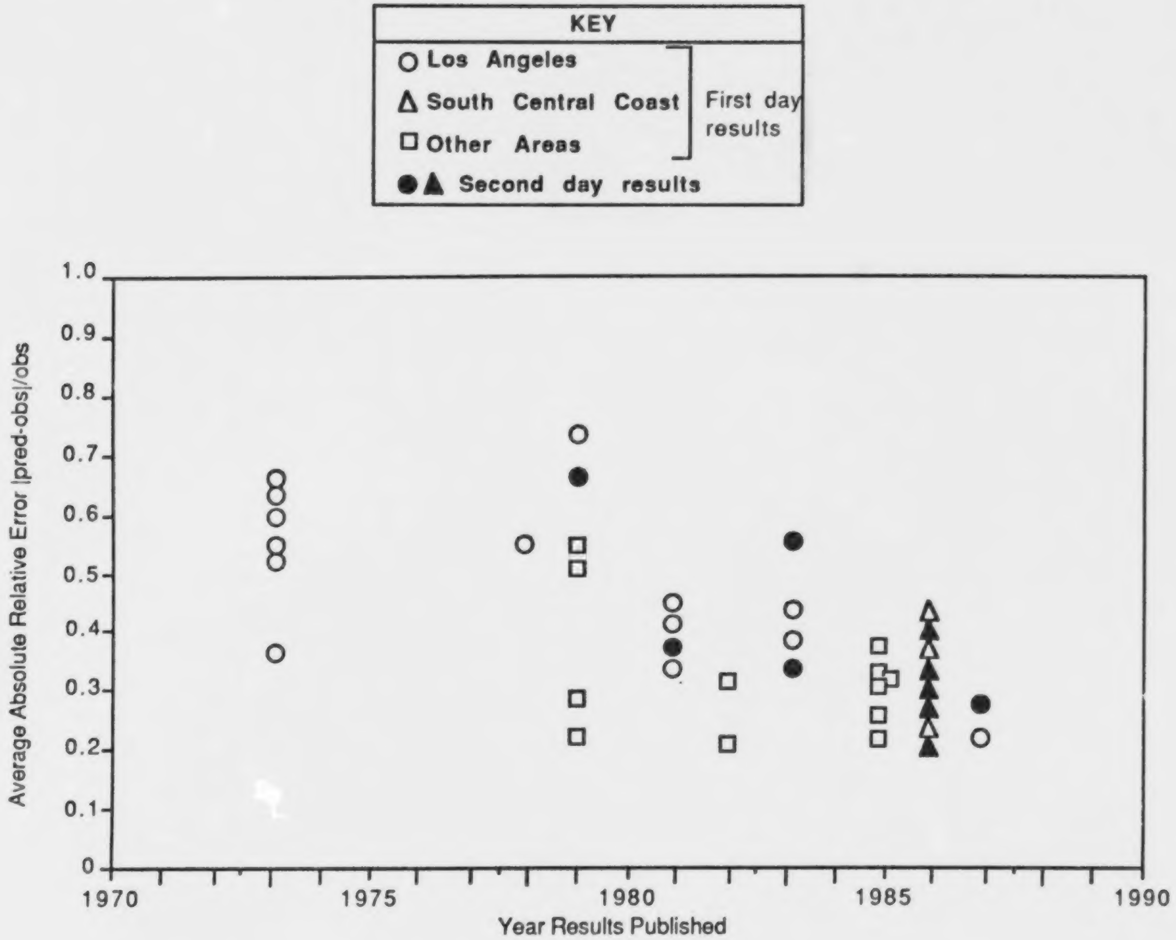


Figure 2. Time history of average error in UAM O_3 results.

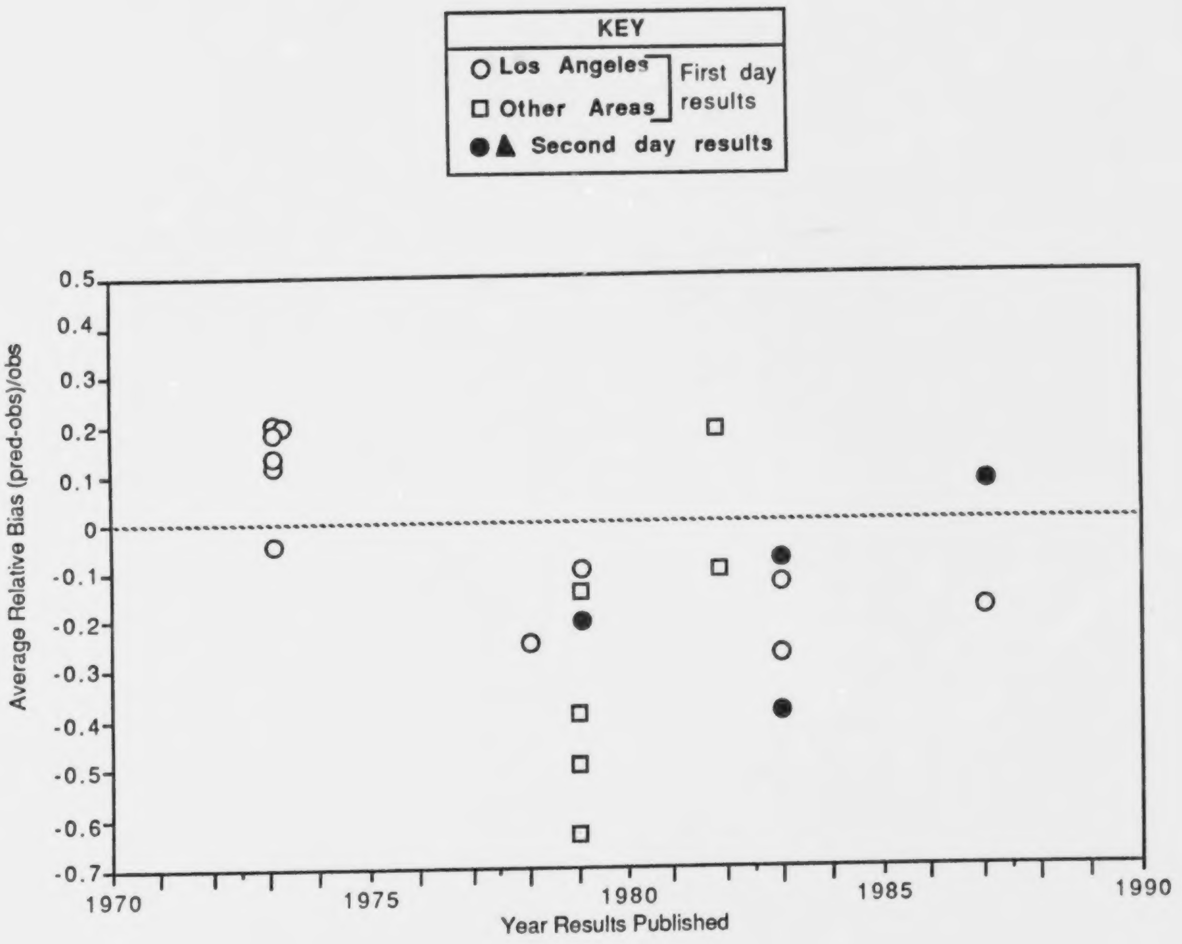


Figure 3. Time history of average bias in UAM NO_2 results.

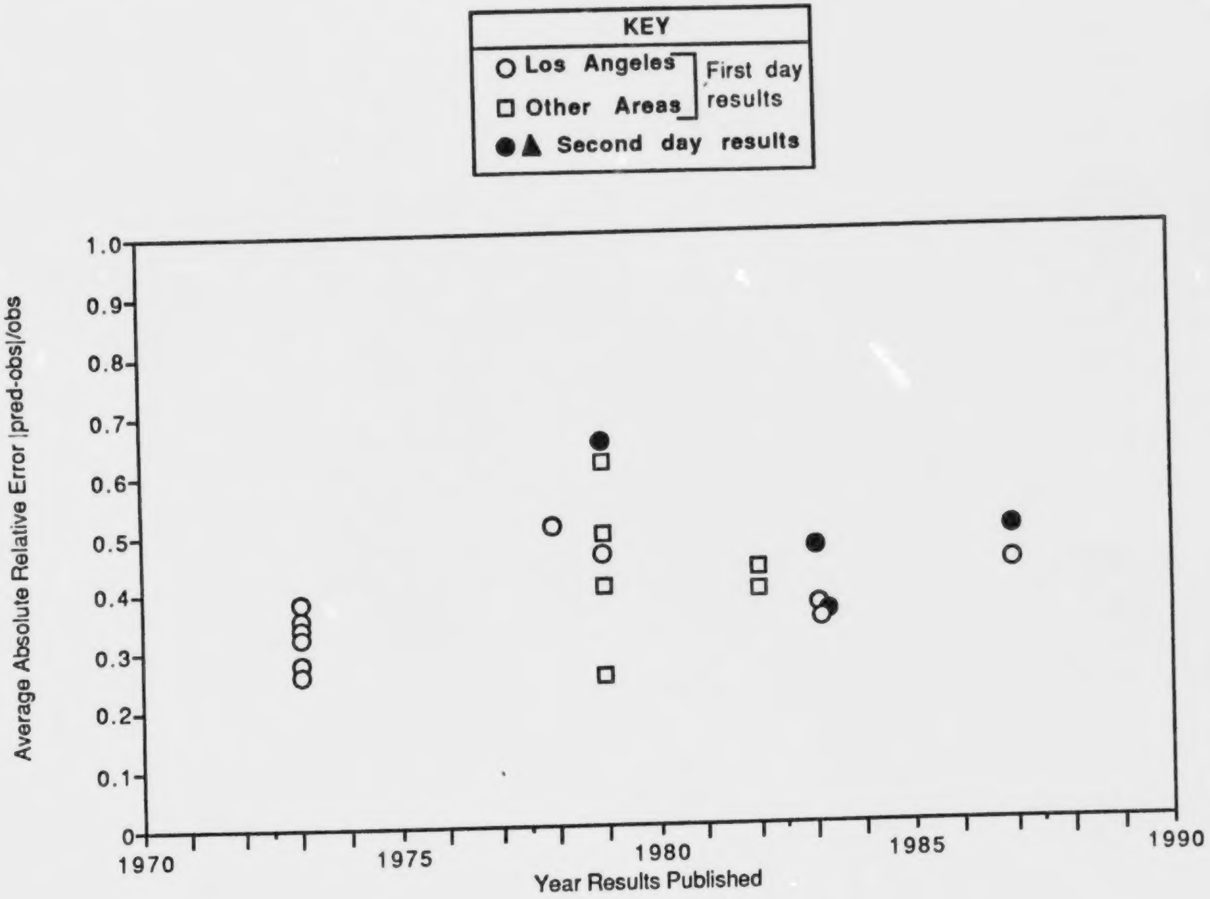


Figure 4. Time history of average error in UAM NO_2 results.

improvements in the absolute error, although there was a significant decrease in the bias for the first simulation day, i.e., 26 June 1974.

- Also, the level of success attained in initial efforts to evaluate performance for a given episode have improved with time. The first attempts to model the 7-8 November 1978 NO_2 episode by Tesche et al. (1983b) yielded a level of performance for ozone somewhat better than that attained in the initial evaluation of the 26-27 June 1974 episode. Furthermore, the recent efforts of Hogo and Yocke (1987) to model the 8-9 August 1984 episode for the first time has yielded results that are as good as any achieved to date.

- Recent applications to Los Angeles, the South Central Coast Air Basin and other inland areas have yielded comparable performance.

- Discrepancies of less than 20 percent have yet to be achieved suggesting a possible limit on the performance of current urban photochemical modeling capability. (Compare with uncertainties in ozone measurements of 15 to 20 percent.)

There are fewer model performance results for NO_2 than for ozone, largely because previous applications of the UAM were intended to study ozone episodes. With the exception of the Los Angeles applications, NO_2 concentrations in the other study areas are rather low, with peak values generally less than 0.10 ppm. Figure 3 summarizes the NO_2 bias results presented in Table 1. From this figure we make the following observations:

- Excluding the initial results published in 1973, the model has tended to underpredict NO_2 concentrations.

- Results for the NO_2 episode of 7-8 November 1978 in Los Angeles are among those that exhibit the least bias.

- Recent simulations seem to exhibit less NO_2 bias than those in the late 1970s and early 1980s, although an insufficient number of results are available to firmly establish this tendency.

The absolute error in NO_2 simulation results is plotted in Figure 4. From this figure we note the following:

- There is no indication of improvement over time in NO_2 prediction accuracy as measured by the overall error.

- Discrepancies between predicted and measured NO_2 concentrations are of the order of 30 to 50 percent of the measured values.

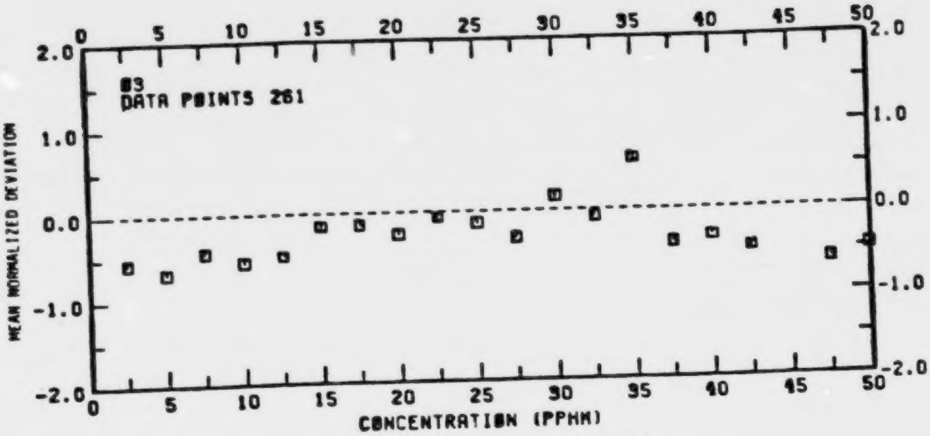
The measures of bias and error presented in Table 1 and Figures 1 through 4 provide only a partial view of the UAM's performance characteristics. For example, the bias and error measures can be calculated using higher threshold concentrations, such as 0.12 or 0.20 ppm for ozone instead of the 0.02 to 0.05 ppm values employed in generating the figures in Table 1. Figures 5 and 6 illustrate the bias and error plotted as a function of measured concentration for the simulation of the 27 June 1974 ozone episode day in Los Angeles. Each figure contains the results of work by Reynolds et al. (1979), Killus et al. (1981), and Tesche et al. (1983a,b).

The results for bias shown in Figure 5 clearly indicate the tendency to underpredict over the entire range of measured concentrations in the initial UAM application to 27 June reported by Reynolds et al. (1979). Subsequent work by Killus et al. (1981) employed an updated version of the Carbon-Bond Mechanism (CBM-II instead of CBM-I) and modifications to the wind fields and initial and boundary concentration inputs. From Figure 5, we note that the overall bias has been reduced, with a tendency to overpredict at lower concentrations and underpredict at higher concentrations. Finally, the third study of 27 June carried out by Tesche et al. (1983a,b) involved further modification of the meteorological inputs, yielding results with little bias for ozone less than 0.35 ppm and a continued tendency to underpredict at higher concentration levels. (Improvements in performance should, of course, be expected during a process of iterative refinement.)

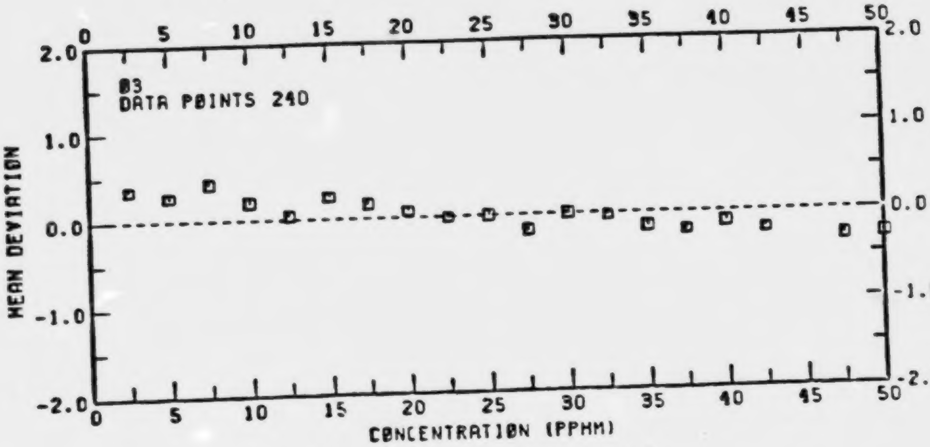
The results for error plotted in Figure 6 indicate that the greatest discrepancies occur for ozone less than 0.10 ppm. The 1981 and 1983 UAM applications exhibit less error for ozone greater than 0.10 ppm than the initial study conducted in 1979.

The bias and error plots of Tesche et al. (1983a) illustrate another aspect of model performance. Note that two sets of results are plotted. In one set of results (designated "weighted average"), the predicted values at each monitoring location are determined using a linear spatial interpolation of the predicted concentrations in the four nearest grid cells. This procedure was also used by Reynolds et al. (1979) and Killus et al. (1981). The second procedure is designated "1 cell away" and involves selecting the predicted concentration that is closest in magnitude to the measured value from the block of nine grid cells immediately surrounding the measurement site. Note that the "1 cell away" procedure yields values for the error that are about half those resulting from use of the "weighted average". The reduction in bias and error is an indication that there is significant spatial variability in the predicted concentration fields.

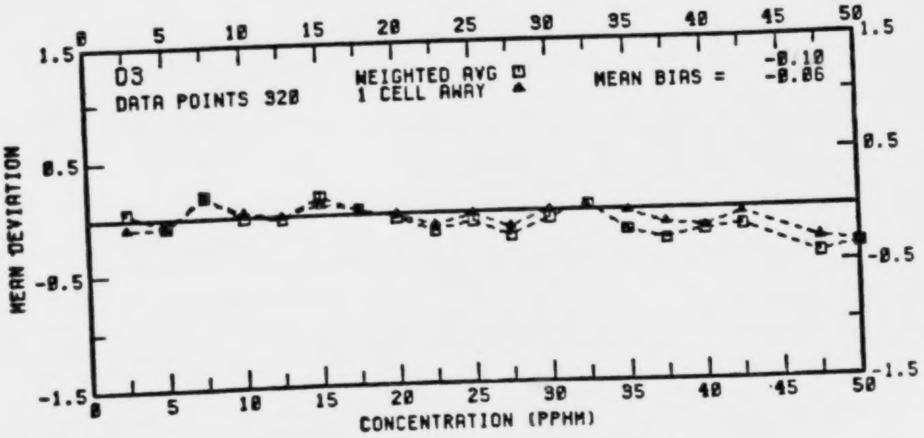
To study the impact of possible uncertainties in model inputs on comparisons between calculated and measured hourly ozone concentrations, Tesche et al. (1983a) examined the results at each monitoring station for the 26-27 June 1974 simulation for each hour between 1000 and 1700 PST to determine the distance between the grid cell in which the monitor was located and the nearest grid cell in which a prediction either equaled or "bracketed" the observed value (i.e., the observed value lay between the value predicted at the cell in which the monitoring station is located and the value at the "bracketing" cell). In all but a very few cases, the station measurement is equaled or bracketed within a distance of seven grid cells. The distribution of distances for the 27 June 1974 results is presented in Figure 7. Note that in over 70 percent of the cases, an ozone prediction comparable to a given measurement can be located within two grid cells, a distance of 5 to 10 km. Some minor discrepancies between predicted and measured values can



(a) Reynolds et al. (1979)

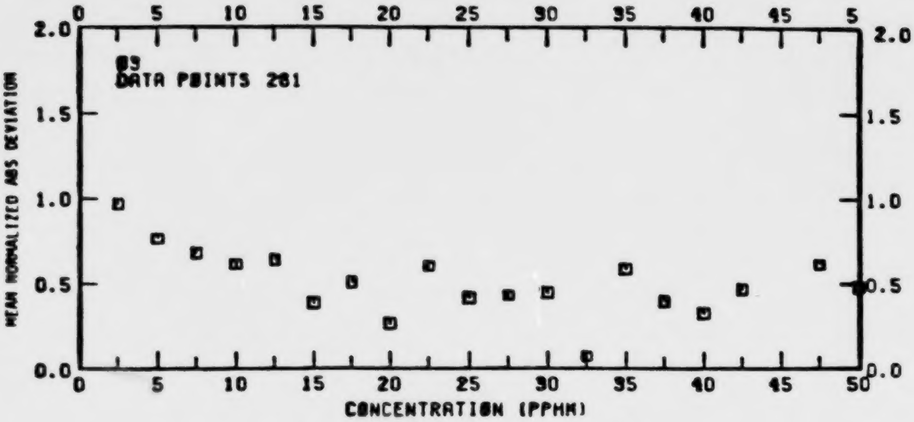


(b) Killus et al. (1981)

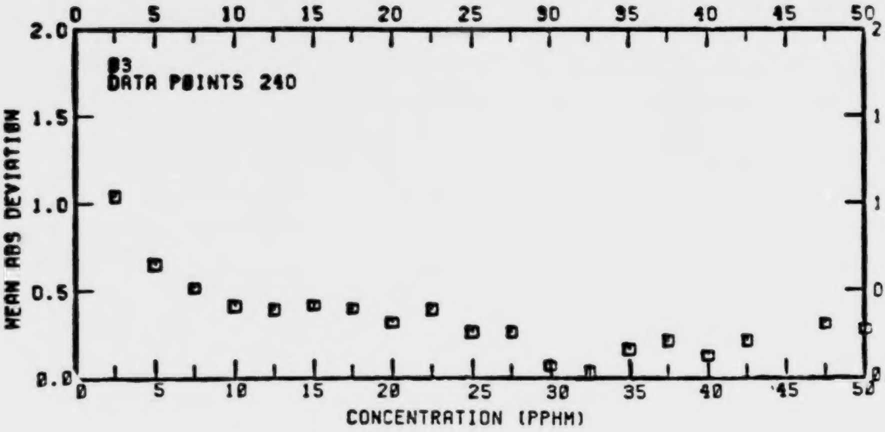


(c) Tesche et al. (1983a)

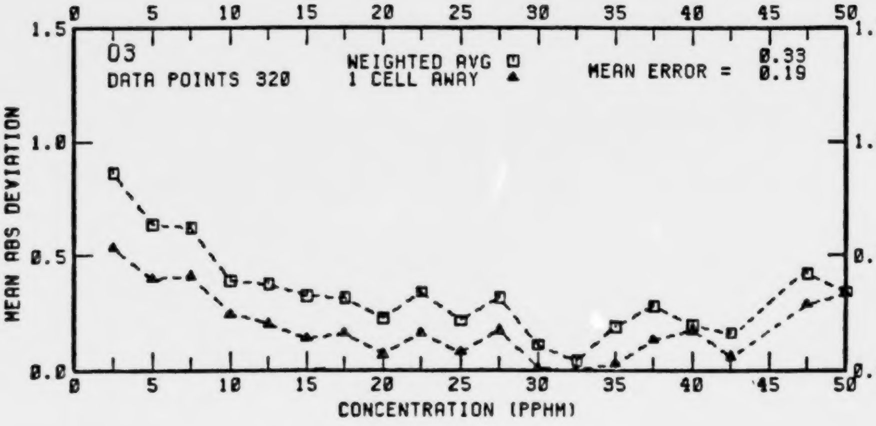
Figure 5. Bias in UAM predictions as a function of measured concentration for three Los Angeles studies--27 June 1974 simulation day.



(a) Reynolds et al. (1979)



(b) Killus et al. (1981)



(c) Tesche et al. (1983a)

Figure 6. Absolute error in UAM predictions as a function of measured concentration for three Los Angeles studies--27 June 1974 simulation day.

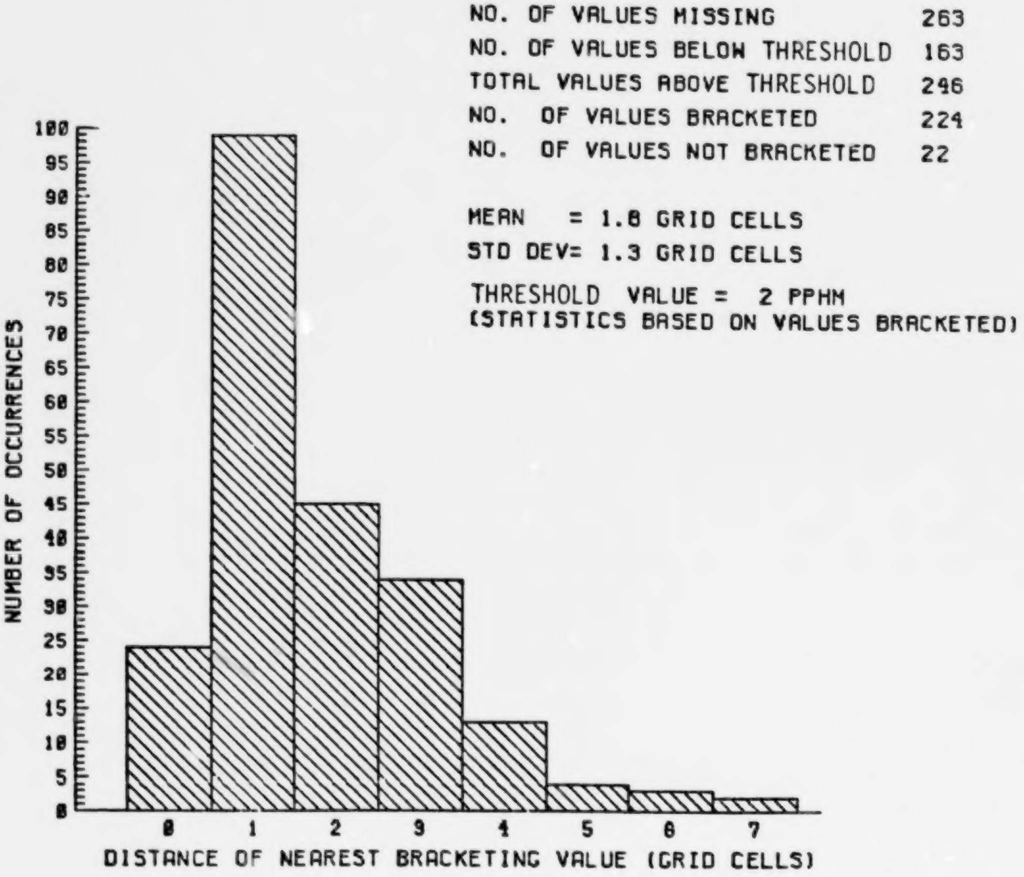


Figure 7. Distribution of distances to nearest bracketed value for O_3 for 27 June 1974. (Source: Tesche et al., 1983a)

be caused by the typical uncertainties in input data, e.g., ~1 m/s in wind speed, or ~20 degrees in wind direction.

Figure 8 illustrates the extent of spatial variability in the vicinity of the Pasadena and Pomona monitoring stations for the simulation of 26-27 June 1974 reported by Tesche et al. (1983a). Note that the observations at the Pasadena station are bracketed by the maximum and minimum predictions within one cell of the station. These plots also illustrate another aspect of model performance. The weighted average predictions (designated "Case D9" in the figure) at Pasadena on 27 June 1974 seem to follow the same general temporal trend as the observations. However, a close examination of these results indicates that the timing is incorrect by one to two hours, yielding significant discrepancies between predicted and measured values, especially during the morning buildup period. Thus, improvements in performance for such results will require that the model closely replicate the timing of the diurnal concentrations at each station.

Another contributor to discrepancies between predicted and observed values is the uncertainty in the measurements. For example, average differences of 15 to 20 percent have been reported for side-by-side comparisons of two independently and routinely operated ozone monitors calibrated by the ultraviolet photometric method (Burton et al., 1976). We note that some of the results presented in Table 1 are approaching this level of measurement uncertainty. For these cases, significant further improvements in performance may be difficult to achieve.

C. A SUMMARY OF FINDINGS

We close this section with the following comments and observations concerning the predictive performance of the UAM:

- The UAM is a deterministic model that attempts to simulate complicated, stochastic atmospheric processes. Limitations in our ability to represent these phenomena in such models, in the number and types of data available with which to specify the required inputs, and in the accuracy of measured concentrations all influence performance levels and the prospects for improvement in performance.

- Studies of the natural variability of the atmosphere suggest that the minimum expected imprecision in model results for ozone is of the order of 10 percent, and possibly larger (Moore and Liu, 1987). Uncertainties in ozone measurements are of the order of 15 to 20 percent. These figures suggest that overall discrepancies between predictions and observations of less than 20 percent may be difficult to achieve.

- A review of UAM performance in studies conducted over the past 15 years indicates that the model has tended to underpredict ozone concentrations in all but the most recent applications. Furthermore, the bias in ozone results has decreased over time,

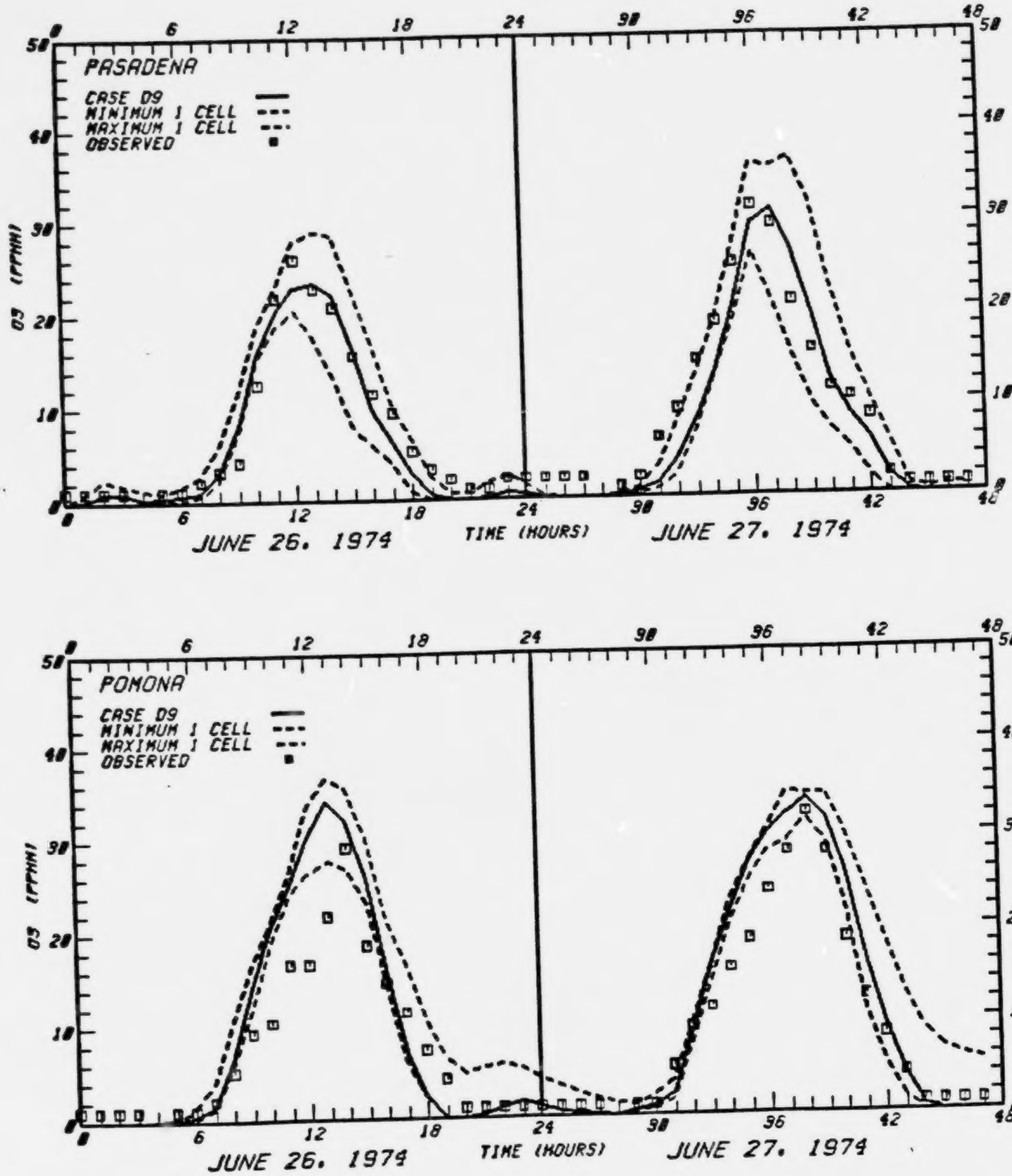


Figure 8. UAM O_3 predictions at Pasadena and Pomona for the 26-27 June 1974 episode. (Source: Tesche et al., 1983a)

presumably because the model now contains better representations of atmospheric processes and model users have become more skilled in preparing inputs.

- The overall error in ozone results has also decreased over the past several years; currently, discrepancies between predicted and observed values are in the range of 20 to 40 percent. The overall error has approached the limit of 20 percent noted above in applications to Los Angeles and other areas as well.

- Improvements in model performance for NO_2 have not paralleled those for ozone over the past 15 years. The overall error in NO_2 results is in the range of 30 to 50 percent.

- When performing sensitivity studies or examining the potential effects of future-year emission control scenarios, care must be exercised in the interpretation of the calculated pollutant concentrations. Since model performance results are likely to exhibit some bias, the predictions for sensitivity and future-year scenarios should be used in a relative manner. That is, the absolute magnitudes of the calculated concentrations should be employed to estimate the expected change in air quality levels resulting from the prescribed change in model inputs.

IV. THE UAM AND POLICY-RELATED STUDIES

In this chapter, we examine the extent to which UAM estimates were utilized as technical input to selected air-quality control decisions. This analysis is based on detailed interviews with participants who were involved in proceedings where UAM estimates were presented as technical input. The broad questions to be answered by these interviews were:

- Was the model acceptable?
- Did its degree of acceptability change over time?
- How was model acceptability related to model performance (as demonstrated by performance evaluations and sensitivity analyses)?

A. POLICY EVENTS

Four applications of the UAM were selected for discussion with interviewees; each application involved a major policy decision pertaining to air quality. Since the policy events were of a varied nature, we will generally refer to them simply as "activities". The four activities were all located in the South Coast Air Basin (SOCAB) of California. By focusing on the SOCAB, we were able to study a six-year period during which UAM estimates were submitted as technical input to air-quality decisions, and during which most of the technical and policy issues were relatively unchanged.

The first activity was the adoption of Rule 1135.1 by the California Air Resources Board (CARB) in 1979, which established emissions requirements for electric utilities. The next activities were the 1979 Air Quality Management Plan (AQMP) and its 1982 revision; both plans were drafted by the South Coast Air Quality Management District (SCAQMD) as part of California's State Implementation Plan (SIP), which was required by the federal Clean Air Act (CAA). The final activity was a set of hearings in 1985 by a subcommittee of the SCAQMD board on the need for additional NO_x controls in the SOCAB.

In addition to these four activities, several others were discussed with interviewees to a lesser extent. One was CARB's 1986 adoption of emission requirements in Kern County. A second concerned model applications in Santa Barbara County: the 1986 Joint Interagency Modeling Study (JIMS) and continuing work undertaken in preparation for hearings on the adoption of emission rules there. Additional applications of the UAM to Tulsa, Philadelphia, St. Louis, and Denver were also discussed with some interviewees. The applications outside the SOCAB were of secondary interest because the technical and policy issues were more varied. We did not exclude the non-SOCAB applications because they raised some significant issues; more importantly, a number of the interviewees were involved in both SOCAB and non-SOCAB applications and chose to offer comments pertaining to the latter as well as the former.

B. SPECIFIC QUESTIONS OF CONCERN

In analyzing the role of the UAM in the selected activities, we wished to answer the following key questions:

1. Were the UAM estimates acceptable as technical input to air-quality control decisions?

2. Did the degree of acceptance of UAM estimates change from 1979 to 1986?

    a. To what extent was modeling information used in each activity (specifically in evaluating control strategies)?

    b. Which other factors played a role?

3. How did performance evaluations, sensitivity analyses, and other factors affect acceptance?

    a. To what extent did performance evaluations lend confidence in model estimates?

    b. Which performance measures most affected assessments of the UAM?

    c. Which measures were desired by participants but not available to them?

    d. To what extent did sensitivity analyses and other factors (e.g., consistency of model estimates with other information) lend confidence in model estimates?

    e. How did particular performance levels translate into rejection, acceptance, or conditional acceptance of model estimates?

    f. Which improvements in performance or sensitivity would have been needed to increase acceptance of the model?

4. What has been and now is the range of opinion regarding the role to be played by photochemical air-quality models?

    a. What is the range of opinion concerning the appropriate uses of photochemical air-quality models as of 1975?

b. Were there significant changes in the accepted use of photochemical models between 1975 and 1986?

c. If changes occurred in the acceptance of photochemical models, were they attributable to perceived improvements in performance?

In posing the preceding questions, it was necessary to define "acceptance". Rather than using an "either-or" definition (acceptable versus unacceptable), we used a subjective definition involving four levels of acceptance:

1. Inclusion of model estimates among technical inputs, but little or no consideration of the estimates.
2. Serious consideration of model estimates.
3. Substantial influence of model estimates.
4. Use of model estimates as one of the principal technical inputs for policy decisions.

By using a four-level definition of acceptance, we sought to accomplish two goals. First, we wished to avoid an unrealistic categorization of the model as either "acceptable" or "unacceptable". Second, we hoped to avoid presuppositions that the model was either "acceptable" or "unacceptable". We used the four-level definition of acceptance to design interview questions that could be answered in terms of degrees.

C. APPROACH

To answer the questions posed in the preceding section, we interviewed the key policy and technical people involved with each of the four activities. We attempted to minimize the time commitments for respondents, while ensuring the utility of the responses. To do so, we conducted telephone interviews.

We first developed a list of individuals who had been involved in one or more of the four primary activities. We selected individuals from five distinct backgrounds:

- Regulatory affiliation--board members,
- Regulatory affiliation--management staff,
- Regulatory affiliation--technical staff,
- Industry affiliation, and
- Scientific peer group (others involved in air-quality modeling).

Individuals with a regulatory affiliation represented the EPA, the CARB, and the SCAQMD. Those with an industry affiliation represented Southern California Edison (SCE) and the Western Oil and Gas Association (WOGA). Peer group representatives were university faculty.

A questionnaire was written and mailed to 38 individuals. Each individual was then contacted by telephone, at which point we requested a telephone interview. The interviews were conducted after each person had had an opportunity to review the questionnaire. The questionnaire (see Appendix C) included both closed- and open-ended questions. We thus obtained both summary statistics and highly useful elaborations on important points.

Although we used a questionnaire to help structure the interviews, this study was not a statistical survey. Respondents were not selected randomly to represent a larger population; rather, we selected respondents on the basis of their expertise. We spoke to a limited number of people who represented a significant portion of the total involved in SOCAB air-quality decisions. We used a telephone-interview technique because the questions that we wished to answer did not require lengthier in-person interviews.

D. SURVEY RESULTS

Of the 38 individuals contacted, 31 agreed to be interviewed.[1] The numbers who participated in each activity, disaggregated by affiliation, are listed in Table 2.

TABLE 2. NUMBER OF SURVEY RESPONDENTS BY ACTIVITY AND AFFILIATION

	Activity						
Affiliation	Rule 1135	1979 AQMP	1982 AQMP	1985 NO_x	Santa Barb.	Kern Cty.	Other
Regulatory: board	4	4	5	3	0	0	0
Regulatory: management	6	7	7	4	4	4	2
Regulatory: technical	2	2	3	3	1	1	0
Industry	3	3	5	5	1	1	0
Scientific peer group	2	1	2	2	0	0	0
Total	17	17	22	17	6	6	2

---

[1] Two individuals who have worked closely together chose to discuss the survey and provide one set of responses. Therefore, 30 separate responses were obtained. In addition, one individual reviewed the responses provided by an associate and wrote to confirm his concurrence with those responses. Two other respondents chose to include colleagues; these two interviews were conducted as conference calls. One respondent replied to a limited number of questions because his involvement in the study activities was minimal.

1. Influence of the UAM

We asked the respondents to rate the influence of UAM modeling results on their assessments of the efficacy of control measures that were under consideration during each activity. The results are listed in Table 3 as numbers of respondents.[2] The influence of the UAM estimates on the respondents' assessments of control measures exhibits modest changes over the period from 1979 to 1982, and a more significant change since 1985. The percentage who ranked the model as having either a substantial impact on their thinking or a more substantial impact than any other factor was 50 percent for Rule 1135 (1979), 69 percent for the 1979 AQMP, 58 percent for the 1982 AQMP, and 78 percent for the 1985 NO_x hearings. The seven individuals who participated in the Kern County hearings (1986) and the three who participated in the Santa Barbara activities (1986) all stated that the model's influence was "substantial" or "more substantial than any other factor".

TABLE 3. INFLUENCE OF UAM ON RESPONDENTS' ASSESSMENTS OF CONTROL MEASURES

	Number of respondents by activity						
Influence	Rule 1135	1979 AQMP	1982 AQMP	1985 NO_x	Santa Barb.	Kern Cty.	Other
Little or none	5	2	3	1	0	0	0
Marginal or moderate	3	2	5	2	0	0	0
Substantial	5	5	6	6	3	3	0
Most substantial	3	4	5	5	0	4	0
Question not applicable	14	14	8	14	24	23	29
Answer missing	0	3	3	2	3	0	1

The respondents also ranked the UAM's influence on the thinking of others, as each respondent perceived that influence. Table 4 lists numbers of respondents by activity and degree of influence.[3] In this table, the increasing influence of the UAM over time should be noted. The percentages of respondents who ranked the model's influence as "substantial" or "more substantial than any other factor" increased from 7 percent for Rule 1135, to 46 percent for the 1979 AQMP, to 53 percent for the 1982 AQMP, to 64 percent for the 1985 NO_x hearings. All individuals who responded to this question for the Kern County or Santa Barbara activities noted the substantial importance of the model. Despite the limited influence of the model from 1979 to 1982, the model results nonetheless affected some individuals' assessments of specific issues, such as the relationship between NO_x emissions and ozone formation.

---

[2] This table corresponds to questions 5a and 5b in the survey (see Appendix C). The total numbers of respondents for each activity differ slightly from the totals listed in Table IV-1. One individual (from the scientific peer group) who was listed as "not participating" in Table IV-1 nonetheless was able to assess the UAM's impact on his thinking. Two individuals participating to a limited extent in some activities could not offer an assessment of the UAM's impact on their thinking because their participation was so limited; these responses were coded as "question not applicable" for questions 5a and 5b.

[3] These results are drawn from questions 6a and 6b. As for Table IV-2, the totals do not add exactly to the totals of Table IV-1.

TABLE 4. INFLUENCE OF UAM ON OTHERS' ASSESSMENTS OF CONTROL MEASURES

	Number of respondents by activity						
Influence	Rule 1135	1979 AQMP	1982 AQMP	1985 NO_x	Santa Barb.	Kern Cty.	Other
Little or none	5	3	4	1	0	0	0
Marginal or moderate	8	4	4	3	0	0	0
Substantial	1	6	8	7	2	2	0
Most substantial	0	0	1	0	0	3	0
Question not applicable	14	14	8	14	24	23	29
Answer missing	2	3	5	5	4	2	1

Twenty-five respondents stated that the model affected their thinking on the nonlinear relationship between NO_x emissions and ozone levels; four stated that it did not. Thus, the model drew attention to the nonlinearity issue, even though it may have had a less than substantial influence on specific policy decisions.

Twenty-one individuals also stated that the model in some way affected their assessment of the number of days that a simulation should be run; eight stated that it did not.

Twenty-one respondents noted that the model led them to recognize the need for a data base specifically designed for use as model inputs; eight did not believe this was an important consideration. Some stated that developing models is a pointless exercise unless a commitment is made to developing a good data base specifically for evaluating and exercising the models.

Twelve respondents stated that the UAM affected their thinking on the utility of grid-based models for assessing the contributions of single sources to ozone formation; fifteen stated that it did not. However, of those answering this question affirmatively, a number meant that the results produced by the UAM had led them to believe that grid-based models could not be used for single-source assessments. Thus, well over half the respondents in fact concluded from the performance of the UAM that grid-based models are incapable of resolving the contributions of single sources.

2. Performance Evaluations

Twenty-three individuals offered an affirmative response to the following question: "Did the existence of performance evaluations, apart from specific details or aspects of the evaluations, affect your assessments of the UAM?" Comments from these twenty-three were typically to the effect that the model would not have been at all credible without the evaluations. Of the four who responded negatively to this question, their additional comments indicated that the performance evaluations were in fact essential, but that it was necessary to go into the details of the evaluations to understand whether or not the evaluations in fact lent credibility to the model.

Table 5 lists numbers of respondents according to activity and their assessment of the performance levels of the model.[4] The respondents' assessments of the model show remarkably little change over time, with most respondents stating that model estimates were good enough to consider, but not good enough to use as the principal basis for formulating control decisions.

TABLE 5. RESPONDENTS' ASSESSMENTS OF UAM PERFORMANCE LEVELS

	Number of respondents by activity				
Performance level	Rule 1135	1979 AQMP	1982 AQMP	1985 NO_x	Other
UAM estimates were:					
Too poor to consider	1	1	0	0	0
Questionable	2	0	1	1	0
Good enough to consider	9	9	11	10	4
Good enough to use for policy decisions	0	0	1	2	0
Question not applicable	16	14	10	13	23
Answer missing	2	6	7	4	3

Twenty-three respondents offered additional comments on model performance. Of these, ten stated that the model was inadequate for quantitative prediction but was adequate for qualitative assessments. These respondents explained that the model was most useful for ranking alternative control strategies: even if the numerical ozone predictions were not very credible, the model's performance was nonetheless sufficiently good that the relative effects of alternative strategies could be tested.

---

[4] From question 3c. For Rule 1135, one respondent stated that performance was between the second and third categories. For the 1979 AQMP and the "Other" category, one and two respondents, respectively, classified performance as between the third and fourth categories. In each of these four cases, the level was coded as the lower of the categories because the respondents did not believe performance was up to the higher levels.

Five respondents mentioned numerical assessments of performance. All were in the range 20 to 40 percent gross error for ozone. One respondent stated that his organization's current criteria for assessing the model's performance include gross error of about 40 percent, bias of 30 percent, and errors on the predicted-to-observed peak of about 30 percent; he would like to see these error limits reduced. Another respondent said that errors of 10 to 15 percent should be the goal. Some respondents also noted that NO_2 prediction was not as good as ozone estimation.

Nine respondents discussed the model's performance in detailed technical terms; they considered these technical aspects to be much more important for assessing the model's performance than the simple summary statistics such as gross error. These technical aspects were:

- NO_2 prediction.
- Understanding correct hydrocarbon:NO_x ratio.
- Variation of performance with distance from emission sources.
- Variation of performance with topography.

Respondents considered these apparent details of major significance because many were concerned that the model could be predicting ozone correctly on average but still be operating incorrectly in some major way, so that predictions of ozone concentration could be very inaccurate in some locations or for some emission scenarios.

Two respondents raised a question concerning the definition of good performance. One stated that performance goals should be established so that you can determine whether or not the model has done well on a performance evaluation. The second pointed out that the adequacy of model performance really depends on how you wish to use the model. The UAM capabilities may be adequate for qualitative but not quantitative assessment at present, which may or may not be acceptable depending upon the desired use of the model.

Two conflicting opinions concerning an alternative approach to modeling were offered. One respondent stated that since 1979 the UAM has been the best assessment tool available, and therefore alternatives such as the Empirical Kinetic Modeling Approach (EKMA) are unsatisfactory. However, a second respondent stated that (as of 1982), EKMA was more immediate, easier to use and probably as good as the more sophisticated models.

3. Performance Measures

We asked respondents to identify the performance measures that they found most useful for judging the UAM's performance. Their responses are listed in Table 6. All measures were utilized to some extent: half the respondents stated either that all measures were important or that no single measure was most important. The comparison of actual to measured spatial and temporal trends was cited most frequently.

TABLE 6. PERFORMANCE MEASURES UTILIZED BY RESPONDENTS

Importance of measure	Number of respondents by measure				
	Gross error	Bias	Peak estimated to peak measured	Temporal and spatial trends	Percentage hours with errors less than 25% or 50%
Not important	10	13	9	5	19
Important	13	12	15	20	9
Very important	5	3	4	3	0
Question not applicable	1	1	1	1	1
Answer missing	1	1	1	1	1

Six of the thirteen respondents who commented on this question discussed the ratio of peak estimated to peak measured ozone. These respondents noted that this measure was important from a regulatory point of view, since the Clean Air Act established standards for maximum hourly concentrations. Six respondents stated that the measures of spatial and temporal trends were most useful from a scientific perspective; one respondent stated that it was most helpful to look at as many graphical representations as possible.

Other than those performance measures listed in Table 6, fourteen respondents identified measures that they had either used or had wished to have available. Of 10 responses concerning other measures used, two respondents mentioned sensitivity analyses (see next section), two noted that their own background and expertise were influential, two identified the model's prediction of the hydrocarbon-to-NO_x ratio, one identified NO_2 performance, one discussed specific uncertainties in the model, one utilized additional information from those with expertise, and one considered the adequacy of performance for a specific application relative to other applications.

Seven respondents identified performance measures that they would have liked to have seen but did not. Two stated that evaluations were needed for species other than ozone (NO_2, hydrocarbons, hydrocarbon-to-NO_x ratio, aldehydes). One requested more detailed discussion of uncertainties, one suggested comparisions of measured concentrations to adjacent grid cells, and one suggested comparison of model errors to the uncertainties associated with other sources of information (e.g., trend analyses). One respondent requested a measure of population exposure and a second suggested comparison of stationary to mobile sources (these measures were not performance measures but rather assessment measures).

4. Sensitivity Analyses

Table 7 shows how interviewees' responded when asked which types of sensitivity analyses, if any, affected their assessments of the accuracy and reliability of UAM estimates. Of twenty-eight responding to this question, 17 indicated that hypothetical emission-reduction scenarios were helpful and 19 affirmed the importance of sensitivity analyses that demonstrated the degree of dependence of the model on uncertainties in emissions or winds, or on initial or boundary concentration conditions. A lesser number, six, indicated that consistency of the model with other information affected their assessment of the model.[5]

TABLE 7. SENSITIVITY ANALYSES UTILIZED BY INTERVIEWEES

	Number of respondents by type of sensitivity analysis			
Response	50 percent emission reductions	Initial or boundary conditions, winds, emissions	Consistency of UAM with other information	Other sensitivity findings
Yes	17	19	6	3
No	11	9	22	25
Question not applicable	2	2	2	2

Comments offered by some of the interviewees emphasized the importance of the sensitivity analyses. Several interviewees stated that the sensitivity analyses were as important as the performance evaluations, and should be considered an integral part of model testing.

Some interviewees noted that there was a need to distinguish between sensitivity analyses that test the model and sensitivity analyses that produce a result. In particular, the sensitivity of the model to 50 percent reductions in NO_x or hydrocarbons can be considered a test of the model (in that ozone concentrations should diminish in response to such large decreases in precursors), but was historically presented as a modeling result.

Some respondents expressed the view that the sensitivities to emission reductions could not in fact be treated as meaningful results unless they represented realistic scenarios.

Many interviewees also noted that one should expect the model to be sensitive to emissions, winds, initial conditions, and boundary concentrations. Therefore, good input data are essential.

---

[5] Other information includes other modeling work, general information on atmospheric chemistry, and analyses of air-quality trends in relation to emission trends.

5. Other Factors Affecting Assessments of Model Accuracy and Reliability

Of those commenting on other factors that affected their assessments of the accuracy and reliability of UAM estimates, four interviewees mentioned their evaluations of the theoretical soundness of the model. Three individuals noted that the honesty of the modelers was a factor. Of these three, one felt that SAI (the model developer) had convinced the scientific and policy communities of their objectivity; two stated that they were made uneasy because the developer testified on behalf of industry clients, although they had no indication of nonobjectivity. Two individuals identified the relationship between modeling and air-quality trend analyses as a factor affecting their assessment of modeling results.

6. Relationship Between Model Performance and Model Influence

Tables 3, 4, and 5 suggest that the UAM's performance, as manifested by performance evaluations, had little to do with acceptance of the model: Tables 3 and 4 indicate increasing influence of UAM estimates as technical inputs to air-quality decisions, but Table 5 indicates little change in the interviewees' assessments of the model's performance over the six-year period.

Some comments offered by the interviewees suggest other factors that may explain increasing influence of the UAM over time. Twenty interviewees offered explanations for the limited influence of the UAM on their thinking. Of these twenty, seven mentioned the need for better input data. Three remarked that EKMA and rollback models more directly addressed what emission reductions were needed during the 1979 to 1982 time period.

Fourteen interviewees commented on factors they thought were crucial in affecting other people's assessments of the UAM. Of these fourteen, six stated that acceptance or rejection of one of the model's key results (that NO_x control was largely ineffective in controlling ozone) was determined by individuals' ositions on the need for emission controls.

Twenty-four respondents identified factors that they believed were more influential than the UAM in affecting the outcomes of specific control decisions that were made. Of these twenty-four, six identified predetermined positions or personal values. Five indicated that political considerations or political compromise were key factors.

In summary, these comments help explain the limited influence of the UAM estimates, particularly during the 1979 - 1982 time period. They do not explain the increase in model acceptance since 1982, however.

7. Interviewees' Views of the Appropriate Uses of Modeling, 1975 - 1986

An explanation for the increased influence of the UAM can be discerned in the interviewees' evaluations of what air-quality models could be used for as of 1975, 1980, or 1986. Of twenty-four comments, only four interviewees indicated that the state-of-the-art in modeling had improved substantially, and one individual said that such changes were minimal. Sixteen respondents stated that since 1980, air-quality models had become useful tools (possibly one of several tools), and most indicated that they were useful in a directional, or qualitative, sense. Some individuals noted that there has been a general recognition that models can deal with the complexities in urban air-quality analysis. A number of interviewees stated that what had changed over this period was not the model itself but their understanding of its capabilities and limitations. These comments are consistent with those described under the section entitled "Performance Assessments", where several interviewees discussed performance in detailed terms and asserted that these apparent details were of major significance for understanding UAM results.

8. Increasing Acceptance of the Model

The respondents were asked which, if any, of a set of factors would have led to a greater influence of the UAM estimates on their assessments of the efficacy of control measures (Table 8). Twenty-one respondents stated that better input data would have led them to place more confidence in the UAM estimates. This identification of the need for better input data is consistent with the previously-discussed affirmative response by twenty-one respondents to a separate question on the need for a data base designed specifically for use as model inputs, as well as separate comments indicating that data limitations affected some respondents' assessments of the reliability of UAM estimates.

TABLE 8. FACTORS THAT WOULD INCREASE ACCEPTANCE OF THE UAM

	Number of respondents by factor			
Response	Better performance evaluations	More consistency with other information	Better input data	More performance evaluations
Yes	10	8	21	16
No	19	21	8	13
Question not applicable	1	1	1	1

The second most frequently cited factor was a greater number of performance evaluations, for a more varied set of meteorological conditions. Several interviewees believed that the model had not been very thoroughly tested because performance evaluations had been conducted for a limited number of days, representing a limited number of meteorological conditions. This need for more evaluations occurred, in part, because no data base for the South Coast Air Basin (or for California) was designed specifically for use as model inputs. Consequently, on only a few days were there data available for running and testing the model.

Two remaining factors were cited less often. Ten respondents felt that they would have been more influenced by UAM estimates if the performance evaluations had been better. Eight respondents stated that they would have been more influenced if the UAM estimates had been more consistent with other technical inputs.[6]

One interpretation of these results is that respondents were less concerned about the particular outcomes of the performance evaluations that were done than they were about known deficiencies in input data. Some interviewees expressed the concern that with much of the input data so poorly measured, the model could be producing reasonable ozone estimates for the wrong reasons. This concern was compounded by having so few performance evaluations available, which represented a limited number of meteorological conditions. A number of respondents were also concerned that the amount of data required to run and test the model satisfactorily could never be obtained.

E. SUMMARY

We summarize here the principal conclusions that emerged from our discussions with interviewees.

1. The extent to which UAM estimates were utilized as technical input to air-quality decisions increased over the period 1979 to 1986.

2. Throughout the period 1979 to 1986, interviewees qualitatively ranked the performance of the UAM as "good enough for estimates to be considered" but not "good enough to provide a high degree of confidence in the model".

3. The increase in acceptance of UAM estimates as technical input was not related to improved performance (or to perceived performance), since the interviewees' qualitative rankings of performance were unchanged over the period 1979 to 1986.

4. The limited initial influence of the model was related to polarization over a controversial finding, a general recognition of the inadequacy of inputs with attendant uncertainty in model results, and an unfulfilled expectation of "acceptable" quantitative performance of the model.

5. The interviewees reported that their understanding of the model's capabilities and limitations improved over the period 1979 to 1986. This improvement in understanding was related to specific aspects of the model, such as the dependence of the NO_x-ozone relationship on the ratio of hydrocarbons to NO_x. The interviewees' understanding of the dependence of model estimates on key technical aspects of model formulation was important to their degree of acceptance of the model.

---

[6] Other technical inputs that were considered during the various activities included other modeling results and empirical analyses of the temporal trends in emissions and air quality.

6. Interviewees now generally consider the model accurate enough for qualitatively assessing the relative effects of alternative control scenarios, but not accurate enough for quantitative evaluation of specific measures. Most respondents stated that they had originally believed that quantitative assessment would be possible, but with improvements in their understanding of the model's capabilities they now concluded that only qualitative comparisons of alternatives was possible.

7. Most interviewees believed that better input data, preferably from a data base specifically designed for use as model inputs, would be necessary to improve the model's reliability and thus its acceptability. They favored the collection of comprehensive data bases designed specifically for supporting model evaluation, with attention given to stressing the model to improve the model's acceptability.

8. Sensitivity analyses were considered as important as performance evaluations for assessing the accuracy and reliability of the model.

V. THE INFLUENCE OF MODEL RESULTS ON POLICY

The primary findings of the preceding chapters are the following:

- The performance of the UAM in predicting ozone concentrations improved with time over the period, 1973 to 1986. Average discrepancies between prediction and observation decreased, as did the spread in average discrepancies among individual studies carried out during the same period. Average bias also decreased; however, the spread in bias among individual studies carried out during the same period remained about constant over time. (See Chapter III.)

  (Note: Predictive performance for NO_2 concentrations was inferior to that for ozone and did not improve with time. However, this finding is not surprising since in most studies much less effort was devoted to studying NO_2 performance and correcting for its deficiencies than was devoted to similar concerns for ozone performance.)

- The acceptance of the UAM results increased with time. (See Chapter IV.)

The purpose of this chapter is to examine the potential significance of these findings, primarily the extent of the relationship between model performance and model acceptance.

Before proceeding with the discussion, however, caveats must be highlighted. First, observations and findings for the most part are subjective. They are derived from the judgments of a limited, albeit significant, group of individuals. Oftentimes, only a fraction -- 1/3 to 1/2 -- of those interviewed felt qualified to respond to particular questions. Sampling of opinion was neither scientific nor statistical. Thus, the findings should be viewed as preliminary, warranting further scrutiny, and the study itself as but a first step in gaining an understanding of the role of modeling in influencing decision making.

A. FACTORS INFLUENCING INCREASED ACCEPTANCE OF THE UAM

While model acceptance and predictive performance both improved (or increased) over the period 1979 to 1986 and thus appear to be correlated, a causal relationship cannot be inferred. In fact, information on time trends in performance were generally unavailable, and individuals interviewed probably were not aware of the nature of the trends. Also, the interviewees' qualitative rankings of performance were unchanged over the period 1979 to 1986, suggesting that the increased level of acceptance of UAM results in policy considerations was attributable to factors other than performance.

Increased acceptance with time appears to derive in part from a combination of (1) limited influence of model results at the beginning of the period under study and (2) greater comfort with modeling as understanding of the models -- their strengths and limitations -- increased. Interviewees indicated that there was an early reluctance to accept modeling results because of the polarization that existed over issues that modeling was to address, the limitations imposed on the value of results due to inadequacies in the supporting data base, and concerns about the quality of performance of the model. With time came considerably improved understanding of the specifics of models and modeling, including the interrelationships among key variables, and of the capabilities and shortcomings of the UAM in general. Thus, greater confidence accompanied judgments, and a greater willingness to take advantage of the strengths of models emerged.

The key ingredients in establishing understanding were:

- dialogue with modelers, in lay terms, about the model -- its formulation, limitations, and attributes.

- discussions of the results of performance evaluations, including at the stage of diagnostic assessment, when possible problems and "cures" were contemplated.

- discussions of sensitivity analyses, when model behavior could be subjectively compared with expectations based on other knowledge and the degree of conformity with those expectations examined. (This is not to say that expectations must be confirmed; sometimes they must be modified. However, the examination of expectation versus model sensitivity is viewed as being of signal importance.)

Other influencing factors favoring model acceptance include:

- the lack of equally attractive tools or options for codifying understanding or synthesizing the elements of complex atmospheric and emissions processes. ("It's the best we have.")

- the model's providing findings that confirm established policy positions. (A somewhat cynical view, but one held by some interviewees, who feel that this factor is important when issues are controversial.)

Those factors impeding model acceptance include:

- obvious or apparent conflicts between the findings of modeling studies and established policy positions.

- inconsistencies between the findings of modeling studies and other analyses (such as correlational and trend analyses).

- concerns about the quality of model performance.

- concerns about a model providing the right answers for the wrong reasons in performance evaluations due to weaknesses in the tests or testing procedure (such as testing for too few, or an inadequate range in, meteorological conditions).

All in all, a number of factors, both technical and policy-related, undoubtedly influenced interviewees' judgments concerning the acceptability of the UAM's results. Thus, while some limited understanding of influences may be inferred, the complexity of the matter precludes determining the relative importance of influences with greater precision. Moreover, the limited nature of the survey itself restricts its use for further clarifying or refining such findings.

B. ASSESSING THE IMPORTANCE OF MODEL "COMPETENCE" IN MODEL ACCEPTANCE

Two primary factors have favored acceptance of the UAM:

- its performance, as gauged by statistical measures, compares favorably with that of other models, and

- its testing, in a formal sense, has been more extensive than that of comparable models.

Interviewees nevertheless indicated that in their view the model's predictions are of value primarily in a relative or qualitative sense. For example, they value the information that an *x%* change in emissions is estimated to produce approximately a *y%* change in peak ozone concentrations, or that one strategy studied appears to be more effective in achieving an objective than rival strategies. However, they have little confidence that a model can be used to reliably estimate (1) the actual peak concentrations (or other, more robust, measures of high concentrations) likely to be result as a consequence of adopting a control strategy, or (2) the extent to which an air quality standard is likely to be approached or met. The current 20% to 40% discrepancies between prediction and observation, the inconsistencies between estimated emissions rates and ambient air quality levels for hydrocarbons, the quality of ambient NO_2 concentration estimates, the uncertainties in emissions models and inventories and other factors provide the basis for this reservation. In effect, the interviewees believe that the average errors associated with model estimates are unacceptably large relative to the magnitude of concentration changes that they are attempting to effect.

Based on these findings, we conclude that "competence" in prediction is clearly a factor in establishing acceptance, as well it must be. Moreover, the quality of UAM's predictions is judged sufficiently good to encourage the decision maker to consider them, in a qualitative or relative sense, in evaluating the merits of alternative control strategies. However, this same quality of prediction also places significant limits on the uses to which decision makers will put these models.

In summary, the UAM model is judged to be valuable, and performance evaluation has been an essential element in establishing confidence in its value. The time required by the regulatory and industrial communities to gain familiarity with and an understanding of the model -- its strengths and shortcomings -- is a key factor in the process of acceptance. The UAM achieves its highest value when it is used in a way that takes advantage of its strengths and recognizes its limitations.

VI. A PERSPECTIVE ON EVALUATING RADM PERFORMANCE: LESSONS FROM THE UAM EXPERIENCE

A primary purpose of conducting this study was to gain perspective about the overall effort surrounding development, evaluation, and application of the RADM. We hoped to learn some lessons from the UAM experience and present them for others to consider in planning and executing RADM-related activities. In this closing section we share thoughts and observations, with the understanding that they are the authors' interpretation of the admittedly limited and subjective findings of Chapter V and prior chapters and, consequently, are subject to discussion, debate, and revision.

A. A "MODEL" OF THE MODELING PROCESS

As discussed briefly in Chapter I, modeling is most appropriately conceived as a more or less continuing process of model development, collection of data, testing, diagnostic analysis, modification and improvement of the model, and repetition of the cycle. This "process model" contrasts with the "product model" -- that of model development to produce a "one shot" or final product.

Modeling is a codification of current knowledge. Knowledge increases in time. When a series of gains in knowledge are made, the model is modified to reflect this new information or understanding. Experience suggests that, for complex atmospheric systems, the time needed to accommodate successive cycles of advancement to the point of diminishing returns is of the order of ten to twenty years. Expectations of policy makers should be based on an understanding of the development process and its associated time frame.

This "model" of the modeling process suggests that a goal of "completing" RADM (i.e., producing a final "product") and applying it in a policy context -- with the model producing reliable results and attaining credibility in the eyes of the scientific and policy communities in the short term -- would be optimistic and possibly naive. About seven years have elapsed since the inception of RADM's development. By 1990 to 1991 the second full round of evaluation will have been completed. Past experience suggests that additional cycles of development and testing will be needed, and that credibility and acceptance may be attained only after perhaps three to four cycles. Thus, if the necessary activities are pursued diligently, the overall time required to achieve model acceptance, based on quality of performance, may be fifteen to twenty years.

While this admittedly subjective time estimate may not be welcomed by policy makers, the news is not all bad. Knowledge gained during the process of development, data acquisition, and diagnostic analysis is generally of considerable value. This knowledge can be used in technical analyses when it becomes available. The findings, in turn, can be of value to policy makers.

Past experience thus suggests that it is unlikely -- although by no means impossible -- that a reliable, credible, generally acceptable model will emerge after a first, or even a second, full round in the developmental cycle. Rather, the "1990 RADM" will encode the best available knowledge at that date. Of course, what policy makers will wish they might have in 1990 is the best available knowledge in the year 2000 or 2005.

Realism requires that expectations -- where they are at odds with past experience -- must be adjusted. One should expect a "best available" model based on "current knowledge" in 1990. One should not expect a reliable, credible, or acceptable model. However, if one should emerge in a shorter time frame, we should certainly wish to take advantage of it.

B. COMPARING THE RADM AND UAM EFFORTS AND EXPERIENCES

Having presented a "model" of modeling, based on the UAM and other experiences, it is appropriate to point out that, while there are significant similarities in both the models and the developmental processes, there are important differences as well. The differences should be considered in developing a perspective on RADM.

1. Regulatory Setting

The RADM is being developed, tested, and applied in a major regulatory context. A great deal of attention is focused on its development and application. A significant level of planning accompanies these efforts. The UAM was developed for the urban scale. Its early development stages were research-oriented in character; no plans were in place to inject UAM into a regulatory setting. Thus, planning was minimal.

This difference in settings would suggest that, with a greater level of attention in planning, the time to acceptance of RADM might be foreshortened relative to that for UAM.

2. Geographical Scale

RADM applies to a spatial scale of subcontinental dimension. UAM applies to a scale of the order of 50 to 300 kilometers. Providing needed information for subcontinental dimensions would appear to dwarf the information requirements for urban scale. However, the spatial resolutions of RADM are coarser than those of UAM, so that overall data requirements for the two models may not differ as much as might first be assumed (although RADM's requirements clearly are the greater).

3. Complexity of Processes

Both UAM and RADM simulate the dynamic behavior of atmospheric and emissions processes that govern pollutant distributions. However, RADM must simulate pollutant uptake into clouds, cloud processes (physical and chemical), and wet scavenging -- a requirement not borne by UAM. Knowledge of these processes is limited; thus, dynamic descriptions of them are still primitive. Field programs designed to support modeling of these processes, such as the Processing of Emissions by Clouds and Precipitation (PRECP) program sponsored by the Department of Energy, hold the key to gaining improved understanding. Such studies -- now being planned or carried out -- will parallel the main monitoring programs that support RADM's performance evaluation. (It is not yet clear if the range of studies will be adequate or the plans for individual studies will be sufficiently encompassing to meet informational needs.) However, the time required to gain a satisfactory understanding of these processes is expected to be a substantial fraction of the time to acceptance of RADM.

4. Nature of Supporting Data Bases

As noted earlier, a primary factor in RADM and UAM development is the availability of data collected in comprehensive field programs to support model performance evaluation. As the main phase of RADM development and initial testing comes to a close, a major monitoring program has been launched to collect data both for providing input to the model and for testing its performance. The RAPS of the 1970s occupies a similar place in the history of UAM performance evaluation; this effort was undertaken at the close of the main phase of model development and provided a primary data base for supporting model evaluation. Unfortunately, little interaction existed between efforts to evaluate and apply the model in California and in St. Louis (RAPS). Thus, UAM (California) development benefitted less from the lessons learned in RAPS than RADM development is likely to benefit from those learned through model exercise using the data from the current ACID-MODES field program.

It is important to distinguish between evaluation of UAM using data obtained in the midwest and in California. Terrain, meteorology, emissions, land-water contrasts, emissions, and air quality differ substantially between these geographical areas. Thus, it is critical that the UAM be tested using California data bases if confidence is to be placed in its use in the state.

Fifteen years elapsed before a field program was first designed and carried out in California to support the evaluation of UAM. UAM testing through the 1970s and the first half of the 1980s was undertaken using routinely collected data bases, some supplemented to an extent with special monitoring. Thus, the potential for achieving adequate performance, and indeed to gain improved understanding of the dynamic processes influencing pollutant concentration distributions, was limited. The availability of a specially designed, comprehensive, supporting data base should, in principle, shorten the time to acceptance of RADM relative to UAM. However, for a variety of reasons, a number of desired measurements could not be included in the ACID-MODES field program. The extent to which this limitation will restrict the potential for adequate evaluation is now under scrutiny.

All in all, considering the various factors discussed here, some favor a longer time to technical acceptance for RADM, some a shorter time. More careful weighing of factors would be unlikely to bring increased precision to forecasts of time to acceptance. What we can say is that (1) field programs tailored to the needs of model development and testing are essential to RADM's attaining acceptance, (2) at least a few cycles of testing and diagnosis are likely to be needed, and (3) improved understanding of cloud processes and wet scavenging are essential to the development of a reliable predictive capability. Our current judgment is that the time to development of a credible, reliable version of RADM that is acceptable to the technical community will not be unlike that required to attain the same objective for UAM under the differing circumstances discussed.

Specifically, allowing for a few cycles of model improvement, data collection, testing, diagnosis, and modification of the model -- and recognizing the ups and downs of funding availability -- a total time to technical acceptance of about fifteen years, and perhaps up to twenty years, might be anticipated.

C. THE ROLE OF PERFORMANCE EVALUATION IN MODEL DEVELOPMENT

As discussed, performance evaluation is an essential step in model development. Interviewees were unanimous in the view that, if confidence in a model is to be achieved, field monitoring programs designed to support evaluation of model performance over a range of conditions must be carried out, and models must be tested using these data. The converse, of course, does not hold; that is, data collection and testing do not assure technical acceptance of a model. A "tested" model may provide "reasonable" estimates of pollutant concentrations or deposition, but it may not do so with sufficient accuracy or reliability to satisfy policy-related needs. Moreover, the time required to carry out a suitable range of tests and to complete two or more cycles of diagnosis, data collection, and model improvement may well be too lengthy to provide "valued information" to the decision maker within his mandated time frame.

What, then, are the implications for RADM? First, it seems clear that the model is unlikely to supply unambiguous, accurate answers to the questions of policy makers in the short term. Rather, the modeling process, as described, will provide gains in knowledge and improved versions of the model which will be of value to the policy maker. Moreover, the modelers become expert and are valued as technical advisers in interpreting findings, despite the limitations of the models themselves. Thus, emphasis should be placed on the process itself, as discussed earlier, rather than on the product, if expectation is to reflect reality.

What is less clear is the extent to which the policy makers' needs will be met by the process itself, or by the model in its current state of development. It is quite possible -- and perhaps likely -- that the time for decision will have passed before model acceptance is achieved. But the experience of UAM and the ozone problem in general suggest that the issue does not vanish because a decision has been made. Ambient air quality standards for ozone (or photochemical oxidants), time scales for attainment, and emissions control requirements have been considered and reconsidered for two decades. If a judgment were rendered by all parties concerned at the first or second decision juncture -- in the 1970s -- that modeling could no longer contribute since decisions were already made, then the UAM would not be available in its current state of development for use in the decision analyses presently being carried out. Thus, experience suggests that, for wide ranging, critical issues, active support for effective model development will not be misplaced, even if there is an apparent mismatch in the time scales of decision makers' needs and model acceptance.

What, then, are the primary requirements for data collection, model testing and development?

- Time. The overall process for development of an acceptable model is time consuming. This must be recognized at the outset to avoid inappropriate or misleading expectations.

- Budget. The overall process of model development, including field programs, is costly. Total costs for the entire endeavor, from initial development to model acceptance, should be realistically estimated at the outset.

- Careful design of all elements of the development and testing process. Such planning assures that monies will be used effectively and that time requirements will be minimized.

- Stressing the model. Stressful testing assures that the flaws in a model will be revealed if the model is indeed deficient. Stressing a model requires that the model be tested over a range of meteorological, emissions, and air quality conditions and that the performance of its individual components be adequately tested as well.

- Successful performance. A model's competency is established through its lack of rejection or failure over a range of tests. Technical acceptance is achieved through (1) the cycle of testing, diagnosis, and improvement and (2) lack of rejection in a series of subsequent tests and applications.

Overall, modeling parallels gains in understanding and codifies them. Models represent the current state of knowledge, provide a basis for asking questions of a "simulator" that integrates the disparate elements comprising the knowledge base, aid in identifying informational needs, and delineate the directions for further inquiry. If current knowledge is inadequate, so will be the model. Models gain acceptance when understanding is adequate.

D. TECHNICAL ACCEPTANCE OF RADM

As discussed, a time frame for technical acceptance of RADM not unlike that for UAM might be contemplated; the time frame for policy acceptance may be foreshortened for the reasons indicated earlier. The main factor likely to inhibit RADM's future technical acceptance is the limited understanding of certain key processes that govern acidic deposition. This incomplete understanding derives primarily from inadequacies, uncertainties, and lacunae in available data bases. The main factor that will promote RADM's technical acceptance is the collection and thoughtful use of a well designed data base in performance evaluation, diagnosis, and further development. If the commitment to obtaining such a data base is compromised, so will be technical acceptance.

The data base now being collected by NAPAP will undoubtedly be valuable in the model development process. It will support gains in understanding, evaluation of model performance, and design of future data bases. However, this data base has very significant limitations -- 24-hour averaging of most species concentrations, limited data collected aloft during intensive sampling periods, no monitoring aloft during storms, and limited spatial resolution. Moreover, the data base will not support testing of individual components of the model -- wind field description, vertical structure (stability, mixing depth, layering), wet and dry deposition processes, emissions rates, and aqueous phase chemistry. Thus, future monitoring activities will be needed to test RADM with sufficient stress to ensure that, if the model is indeed flawed, its inadequacies will be revealed. With these needs in mind, the time scale estimated for developing a reliable model is about fifteen to twenty years (or through about the year 2000), assuming that an adequate budget is made available. Thus, the model is unlikely to provide quantitative predictions of sufficient accuracy to satisfy policy needs in the next few years. Rather, RADM's primary value to the policy community in the intermediate term will lie in its integral role in the cycle of learning and advancing understanding, and in educating modelers who provide counsel to decision makers.

E. THE ACCEPTANCE OF RADM IN POLICY ANALYSIS

Interviewees subjectively judged that UAM provides findings of value in policy analysis. Moreover, their degree of acceptance of findings increased with time. However, they did not feel that they could depend on these findings to guide policy, even after nearly two decades of model development. Quite simply, the uncertainties that attend the use of UAM are significant when compared with the changes in concentrations that are estimated as a response to contemplated changes in emissions.

In considering the time frame for acceptance of RADM, we must distinguish between acceptance by the technical and policy communities. Much of the preceding discussion focuses on aspects of efforts to achieve technical acceptance, such as meeting informational needs and the adhering to the cyclical process for testing and improvement. Acceptance of a model by the policy community is influenced by technical arguments and other factors as well. These factors include (1) degree of demand for the model by decision makers and by managers who fund model development and application and (2) availability and attractiveness of alternative models.

In the case of UAM, demand for the model during the first ten years of its development and application were relatively low. This was attributable to (1) the relative novelty of grid-based models and thus the lack of familiarity of the policy community with them; (2) the availability and relative simplicity of EKMA; and (3) the geographical imbalance in demand for the model at the next stage of development (about 1980) - i.e., the strong interest in California versus the mild interest or disinterest throughout the rest of the country.

In contrast, acidic deposition received intensive attention more than a decade after urban ozone was the focus of regulatory action. Familiarity with models and model use was much higher in the early 1980s, and the pros and cons of employing simpler versus more sophisticated models were clearer in the minds of policy makers and managers, based on the experiences of the past. Moreover, a strong national interest existed in addressing the acidic deposition issue; Congress, in fact, was and is monitoring progress in development of RADM. RADM is a centerpiece of the National Acid Deposition Assessment Program (NAPAP). Thus, the policy community badly wants and needs a sound model for use in analysis. This "mindset" could be conducive to promoting acceptance of an "adequate" or "best available" model more rapidly than might occur (or be justified) based on scientific assessment alone.

A complicating factor in policy acceptance is what might be termed "the involvement syndrome". History suggests that model acceptance is clearly enhanced if interested parties are participants in or financial supporters of the development, testing, and/or application of the model, or if they have been involved in a substantive monitoring and/or review capacity. Nonparticipants and policy or technical representatives of groups, agencies, or interests that might hold different views or have different missions than the supporting agency or group tend to be more skeptical and less accepting of the model. Thus, the model can become the focus of differences, which in turn heightens questioning and delays acceptance. While the need to lessen this factor in judging model acceptability is clearly desirable, it is likely to be diminished only after the interested parties find improved processes for interacting in the development, testing, and application of models and in judging model acceptability. The extent to which "the involvement syndrome" influences time to acceptance of RADM is clearly difficult to gauge; however, it may well be an important contributor to the uncertainties attending any estimate of time frame for policy acceptance.

A number of issues have been raised in this discussion. A number of speculations have been made. Analysis of historical events and policy issues is clearly subjective. Yet, from our perspective further analysis and investigation is warranted. The art and science of modeling is likely to benefit from a fuller understanding of past practice and its impacts on the technical and policy communities.

REFERENCES

Anderson, G. E., S. R. Hayes, M. J. Hillyer, J. P. Killus, and P. V. Mundkur. **1977**. Air Quality in the Denver Metropolitan Region: 1974-2000. Systems Applications, Inc., San Rafael, CA. SAI 77-222; EPA-909/1-77-002.

Bencala, K. E., and J. H. Seinfeld. **1979**. An air quality model performance assessment package. *Atmos. Environ.*, **13**:1181-1185.

Builtjes, P. J. H., and S. D. Reynolds. **1983**. Modeling the effects of emission controls in the Netherlands. *Environ. Inter.*, **9**:573-580.

Burton, C. S., P. J. Bekowies, R. I. Pollack, and P. Connell. **1976**. Oxidant/Ozone Ambient Measurement Methods: An Assessment and Evaluation. Systems Applications, Inc., San Rafael, CA. SAI 76-111R.

CARB. **1985**. Kern County Ozone Modeling for SIP Update. California Air Resources Board, Technical Support Division, Sacramento, CA.

Colorado Department of Health. **1977**. Analysis of Air Quality in the Denver Air Quality Maintenance Area. Air Pollution Control Division.

Colorado Division of Highways. **1976a**. Air Quality Assessment Statement. Joint Regional Planning Program. CY-1976.

Colorado Division of Highways. **1976b**. I-470 Detailed Assessment Report.

Deardorff, J. W. **1972**. Numerical investigation of neutral and unstable planetary boundary layers. *J. Atmos. Sciences*, **29**:91-115.

Dennis, R. L., M. W. Downton, and R. S. Keil. **1983**. Evaluation of Performance Measures for an Urban Photochemical Model. U.S. Environmental Protection Agency. EPA-450/4-83-021.

Donnelly, D. E. **1978**. Oxidant Model Applications: Denver. 57th Annual Transportation Research Board Meeting, Washington, D.C.

Eschenroeder, A. Q., J. R. Martinez, and R. A. Nordsieck. **1973**. Evaluation of a Diffusion Model for Photochemical Smog Simulation. U.S. Environmental Protection Agency. EPA-R4-73-012a.

Fox, D. G. **1981**. Judging air quality model performance. *Bull. Amer. Meteor. Soc.*, **62**:599-609.

Haney, J. L., and T. N. Braverman. **1985**. Evaluation and Application of the Urban Airshed Model in the Philadelphia Air Quality Control Region. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-450/4-85-003.

Haney, J. L., D. R. Souten, T. W. Tesche, L. R. Chinkin, H. Hogo, and M. C. Dudik. **1986**. Evaluation and Application of the PARIS Photochemical Model in the South Central Coast Air Basin, Volume I. Systems Applications, Inc., San Rafael, CA. SYSAPP-86/065.

Hayes, S. R. **1979**. Performance Measures and Standards for Air Quality Simulation Models. Systems Applications, Inc., San Rafael, CA. SAI 78-93R; EPA-450/4-79-032.

Hecht, T. A. **1972**. Further Validation of a Generalized Mechanism Suitable for Describing Atmospheric Photochemical Reactions. Systems Applications, Inc., San Rafael, CA. 72-SAI-26.

Hecht, T. A., P. M. Roth, and J. H. Seinfeld. **1973**. Mathematical Simulation of Photochemical Reactions: Model Development, Validation, and Applications. Systems Applications, Inc., San Rafael, CA. R73-28.

Hecht, T. A., and J. H. Seinfeld. **1972**. Development and Validation of a Generalized Mechanism for Photochemical Smog. *Environ. Sci. Technol.*, **6**:47-57.

Hecht, A., J. H. Seinfeld, and M. C. Dodge. **1974**. Further Development of Generalized Kinetic Mechanism for Photochemical Smog. *Environ. Sci. Tech.*, **8**:327ff.

Hoel, T. G. **1962**. *Introduction to Mathematical Statistics*. Wiley, New York, NY.

Hogo, H., and M. A. Yocke. **1987**. Evaluation of a Photochemical Air Quality Model with Extensions to Calculate Aerosol Dynamics and Visibility. Systems Applications, Inc., San Rafael, CA. SYSAPP-87/017.

Jerskey, T. N., and J. H. Seinfeld. **1976**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Volume IV: Examination of the Feasibility of Modeling Photochemical Aerosol Dynamics. Systems Applications, Inc., San Rafael, CA. SAI 75/26R; EPA-600/4-76-016d.

Jerskey, T. N., J. H. Seinfeld, F. Gelbar, and L. E. Reid. **1977**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Volume VII: Mathematical Modeling of Urban Aerosol Dynamics. Systems Applications, Inc., San Rafael, CA. SAI 77/144R.

Killus, J. P., J. P. Myer, D. R. Durran, G. E. Anderson, T. N. Jerskey, S. D. Reynolds, and J. Ames. **1977**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Vol. V: Refinements in Numerical Analysis, Transport, Chemistry, and Pollutant Removal. Systems Applications, Inc., San Rafael, CA. SAI 77/142R.

Killus, J. P., W. R. Oliver, P. D. Gutfreund, J. E. Langstaff, T. W. Tesche, H. J. Su. **1981**. Simulation of Impacts of Nitrogen Oxide Control Strategies Under Oxidant Episode Conditions. Systems Applications, Inc., San Rafael, CA. SYSAPP-81/165.

Lamb, R. G. **1976**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Volume III: Modeling of Microscale Phenomena. Systems Applications, Inc., San Rafael, CA. SAI 75/25R; EPA-600/4-76-016c.

Lamb, R. G., W. H. Chen, and J. H. Seinfeld. **1975**. Numerico-empirical analyses of atmospheric diffusion theories. *J. Atmos. Sci.*, **32**:1794-1807.

Lamb, R. G., W. R. Shu, D. R. Durran, J. H. Seinfeld, and L. E. Reid. **1977**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Vol. VI: Further Studies in the Modeling of Microscale Phenomena. Systems Applications, Inc., San Rafael, CA. SAI 77/143R.

Layland, D. E., and H. S. Cole. **1983**. A Review of Recent Applications of the SAI Urban Airshed Model. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-450/4-84-004.

Liu, M. K., and J. H. Seinfeld. **1975**. On the validity of grid and trajectory models of urban air pollution. *Atmos. Environ.*, **9**:555-574.

Liu, M. K., D. C. Whitney, and P. M. Roth. **1976b**. Effects of atmospheric parameters on the concentration of photochemical pollutants. *J. Appl. Meteor.*, **15**:829-835.

Liu, M. K., D. C. Whitney, J. H. Seinfeld, and P. M. Roth. **1976a**. Continued Research in Mesoscale Air Pollution Simulation Modeling--Vol. I: Analysis of Model Validity and Sensitivity and Assessment of Prior Evaluation Studies. Systems Applications, Inc., San Rafael, CA. SAI 75/23R; EPA-600/4-76-016a.

Moore, G. E., and M. K. Liu. **1987**. Time Series and Spectral Analyses of Concentration and Meteorological Data from the Kincaid and Bull Run Power Plants. Systems Applications, Inc., San Rafael, CA. SYSAPP-87/014.

Oliver, W. R., H. Hogo, and P. Saxena. **1983**. Assessment of NO_x Emission Control Requirements in the California South Coast Air Basin, Volume IV: Appendix C, Emission Inventory Review and Development. Systems Applications, Inc., San Rafael, CA. SYSAPP-83/053.

Pacific Environmental Services. **1973**. Controlled Evaluation on the Reactive Environmental Simulation Model (REM), Volume I. U.S. Environmental Protection Agency. EPA-R4-73-013a.

Reynolds, S. D. **1977**. The Systems Applications, Inc. Urban Airshed Model: An Overview of Recent Developmental Work. In: *International Conference on Photochemical Oxidant Pollution and Its Control*. U.S. Environmental Protection Agency. EPA-600/3-77-001b.

Reynolds, S. D., J. Ames, T. A. Hecht, J. P. Meyer, D. C. Whitney, and M. A. Yocke. **1976**. Continued Research in Mesoscale Air Pollution Simulation Modeling, Volume II: Refinements in the Treatments of Chemistry, Meteorology, and Numerical Integration Procedures. Systems Applications, Inc., San Rafael, CA. SAI 75/24R; EPA-600/4-76-016b.

Reynolds, S. D., J. Ames, G. W. Lundberg, and P. V. Mundkur. **1978**. Application of the SAI Airshed Model to the Evaluation of Alternative Population Growth Forecasts for the South Coast Air Basin. Systems Applications, Inc., San Rafael, CA. SAI 78-124.

Reynolds, S. D., H. Hogo, W. R. Oliver, and L. E. Reid. **1982**. Application of the SAI Airshed Model to the Tulsa Metropolitan Area. Systems Applications, Inc., San Rafael, CA. SYSAPP-82/004.

Reynolds, S. D., M. K. Liu, T. A. Hecht, P. M. Roth, and J. H. Seinfeld. **1973a**. Further Development and Validation of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants. Systems Applications, Inc., San Rafael, CA. SAI 73/19.

Reynolds, S. D., M. K. Liu, T. A. Hecht, P. M. Roth, and J. H. Seinfeld. **1974**. Mathematical modeling of photochemical air pollution--III. Evaluation of the model. *Atmos. Environ.*, **8**:563-596.

Reynolds, S. D., L. E. Reid, and M. J. Hillyer. **1979**. Photochemical Modeling of Transportation/Control Strategies--Vol. I. Model Development, Performance Evaluation, and Strategy Assessment. Systems Applications, Inc., San Rafael, CA. SAI 79/37R.

Reynolds, S. D., and J. H. Seinfeld. **1975**. Interim evaluation of strategies for meeting ambient air quality standards for photochemical oxidants. *Environ. Sci. Technol.*, **9**:433-447.

Reynolds, S. D., J. H. Seinfeld, and P. M. Roth. **1973b**. Mathematical modeling of photochemical air pollution--I. Formulation of the model. *Atmos. Environ.*, **7**:1033-1061.

Ross, D. **1975**. Staff Report on the Preparation of Data Files for the SAI Model. California Department of Transportation, District 7, Environmental Investigations Section.

Roth, P. M., S. D. Reynolds, P. J. W. Roberts, and J. H. Seinfeld. **1971**. Development of a Simulation Model for Estimating Ground-Level Concentrations of Photochemical Pollutants. Systems Applications, Inc., San Rafael, CA. SAI 71/21.

Roth, P. M., S. D. Reynolds, T. W. Tesche, P. D. Gutfreund, and C. Seigneur. **1984**. An appraisal of emissions control requirements in the California South Coast Air Basin. *Environ. Inter.*, **9**:549-571.

Roth, P. M., P. J. W. Roberts, M. K. Liu, S. D. Reynolds, and J. H. Seinfeld. **1974**. Mathematical modeling of photochemical air pollution--II: A model and inventory of pollutant emissions. *Atmos. Environ.*, **8**:97-130.

Roth, P. M., C. Seigneur, S. D. Reynolds, and T. W. Tesche. **1983**. Assessment of NO_x Emission Control Requirements in the California South Coast Air Basin, Volume IV: Results, Findings, and Implications for Control Requirements Based on Short-term Modeling Studies. Systems Applications, Inc., San Rafael, CA. SYSAPP-83/157.

Schiermeier, F. A. **1978**. Air monitoring milestones: RAPS' field measurements are in. *Environ. Sci. Technol.*, **12**:644-651.

Seigneur, C., S. D. Reynolds, P. M. Roth, J. L. Haney, H. Hogo, M. C. Dudik, W. R. Oliver, and T. W. Tesche. **1983b**. Assessment of NO_x Emission Control Requirements in the California South Coast Air Basin, Volume IV: Appendix D, Model Simulations of the Effect of ROG and NO_x Emission Reductions on Short-term O_3 and NO_2 Ambient Concentrations. Systems Applications, Inc., San Rafael, CA. SYSAPP-83/159.

Seigneur, C., T. W. Tesche, P. M. Roth, and M. K. Liu. **1983a**. On the treatment of point source emissions in urban air quality modeling. *Atmos. Environ.*, **17**:1655-1676.

Souten, D. R., T. W. Tesche, and W. R. Oliver. **1981**. Use of the Airshed Grid Model to Evaluate the Air Quality Impacts of Alternative Air Pollution Control Policies on the South Coast Air Basin. Systems Applications, Inc, San Rafael, CA. SAI 80-220.

Stern, R., and B. Scherer. **1984**. Anwendung eines komplexen Ausbreitungsmodells zur Simulation einer photochemischen Smogepisode. Umweltbundesamt, Berlin. 104 02 551/01.

Tesche, T. W., and C. S. Burton. **1978**. Simulated Impact of Alternative Emissions Control Strategies on Photochemical Oxidants in Los Angeles. Systems Applications, Inc., San Rafael, CA. SAI 78/22R.

Tesche, T. W., C. S. Burton, and V. A. Mirabella. **1979**. Recent Verification Studies with the SAI Urban Airshed Model in the South Coast Air Basin. In: *Proc. of Fourth Symposium on Turbulence, Diffusion, and Air Pollution*, 15-18 January 1979, Reno, NV.

Tesche, T. W., W. R. Oliver, H. Hogo, P. Saxena, and J. L. Haney. **1983a**. Assessment of NO_x Emission Control Requirements in the California South Coast Air Basin. Volume IV: Appendix A--Performance Evaluation of the Systems Applications Airshed Model for the 26-27 June 1974 O_3 Episode in the South Coast Air Basin. Systems Applications, Inc., San Rafael, CA. SYSAPP-83/037.

Tesche, T. W., W. R. Oliver, H. Hogo, P. Saxena, and J. L. Haney. **1983b**. Assessment of NO_x Emission Control Requirements in the California South Coast Air Basin, Volume IV: Appendix B--Performance Evaluation of the Systems Applications Airshed Model for the 7-8 November 1978 NO_2 Episode in the South Coast Air Basin. Systems Applications, Inc., San Rafael, CA. SYSAPP-83/038.

Tesche, T. W., W. R. Oliver, D. R. Souten, H. Hogo, and J. L. Haney. **1981a**. Airshed Model Sensitivity to Emissions Inventory Perturbations--26-27 June 1974 South Coast Air Basin Oxidant Episode. Systems Applications, Inc., San Rafael, CA. SYSAPP-81/054.

Tesche, T. W., C. Seigneur, L. E. Reid, P. M. Roth, W. R. Oliver, and J. C. Cassmassi. **1981b**. Sensitivity of Complex Photochemical Model Estimates to Detail in Input Information. Systems Applications, Inc., San Rafael, CA. SYSAPP-81/004, 005, 006.

Tesche, T. W., and D. C. Whitney. **1975**. Treatment of the Input Data for the Expanded Air Pollution Simulation Program Modeling Region. Quarterly Technical Program Narrative, Systems Applications, Inc., San Rafael, California.

Whitten, G. Z., and H. Hogo. **1977**. Mathematical Modeling of Simulated Photochemical Smog. U.S. Environmental Protection Agency. EPA-600/3-77-011.

Whitten, G. Z., T. C. Meyers, C. Daly, L. R. Chinkin, S. D. Reynolds, N. M. Yonkow, B. Austin. **1985**. Application of the Urban Airshed Model to Kern County. Systems Applications, Inc., San Rafael, CA. SYSAPP-85/200.

Yocke, M. A., R. E. Morris, H. Hogo, L. R. Chinkin, and L. A. Mahoney. **1985**. Analysis of the Air Quality Impacts of the San Miguel Project. Systems Applications, Inc., San Rafael, CA. SYSAPP-85/127.

APPENDIX A

TECHNICAL FEATURES OF THE SYSTEMS APPLICATIONS, INC. URBAN AIRSHED MODEL

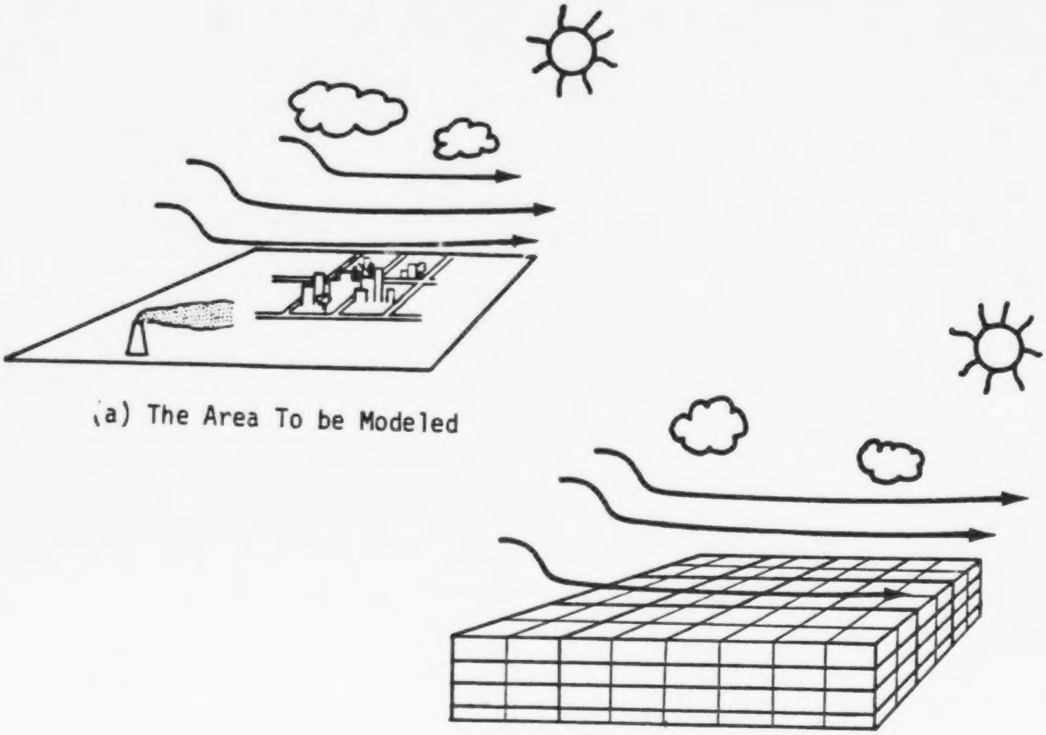
This appendix provides an overview of the technical features contained in the Systems Applications, Inc. Urban Airshed Model (UAM). It is important to note that many of these features have changed over time as better representations of atmospheric processes have been developed. This discussion will identify the important components of the UAM and highlight some of the model improvements that have been implemented over the last 15 years.

The original UAM developed by Roth et al. (1971) and Reynolds et al. (1973a) simulated the dynamic behavior of six pollutants: reactive and unreactive hydrocarbons, NO, NO_2, ozone, and CO. Currently, the model treats up to 20 pollutants. When photochemical simulations are carried out, about 10 to 16 species must be included depending on the formulation of the chemical kinetic mechanism included in the model. In addition to simulating ozone, NO, NO_2, and CO, a number of organic and nitrogen-bearing compounds must also be included. Reactive organic compounds may include paraffins, olefins, aromatics, and aldehydes, while nitrogen-bearing species may include nitrous acid (HONO), nitric acid (HNO_3), and peroxyactyl nitrate (PAN). In addition, the model can provide predictions of SO_2 and aerosol. The user can exercise several options concerning the species to be simulated. However, when photochemical simulations are carried out, a minimum set of organic compounds and nitrogen-bearing species must be included to enable chemical reaction rates to be determined.

As illustrated in Figure A-1, simulations are performed on a three-dimensional grid selected by the user to cover the region of interest. Figure A-1 also shows the various atmospheric processes treated in each grid cell. Early formulations of the model simulated pollutant dispersion in the region between the ground and the base of an elevated inversion layer. More recent versions include provisions enabling the user to model transport and dispersion within both the mixed and inversion layers. The computer codes have been structured to allow inclusion of up to 10 vertical layers of cells and any number of cells in the horizontal directions.

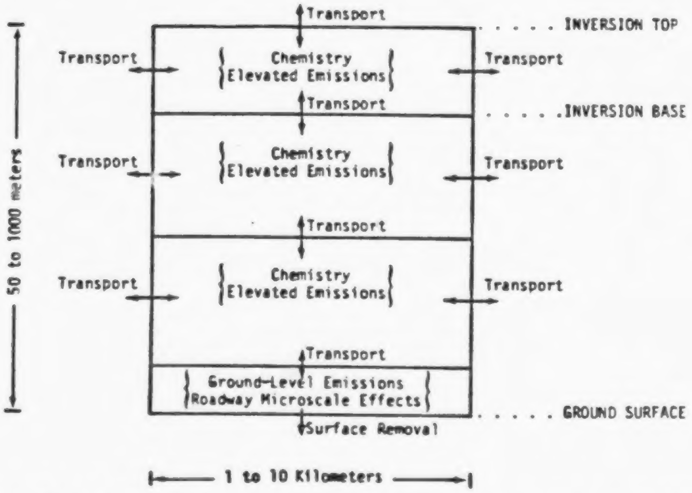
TREATMENT OF ATMOSPHERIC CHEMISTRY

Ozone is not emitted from sources; rather, it is formed by chemical reactions in the atmosphere. Consequently, to calculate ozone concentrations, the UAM must contain a kinetic mechanism, that is, a group of chemical reactions and rate constants intended to represent the ozone formation process. Because of computing time limitations, it is not possible to include explicitly each chemical species of interest; thus,



(a) The Area To be Modeled

(b) Specification of the Grid


Transport
INVERSION TOP
Transport
Chemistry
Elevated Emissions
Transport
Transport
INVERSION BASE
Transport
Chemistry
Elevated Emissions
Transport
Transport
Transport
Chemistry
Elevated Emissions
Transport
Transport
Ground-Level Emissions
Roadway Microscale Effects
GROUND SURFACE
Surface Removal
50 to 1000 meters
1 to 10 Kilometers

(c) Atmospheric Processes Treated in a Column of Grid Cells

Figure A-1. Schematic illustration of the grid used and treatment of atmospheric processes in the SAI Urban Airshed Model.

kinetic mechanisms suitable for use in the model treat organic compounds in groups, often on the basis of the reactive functional groups they contain. The first mechanism included in the UAM was a 15-step scheme developed by Hecht and Seinfeld (1972) and required the simulation of ozone, NO, NO_2, and one lumped reactive hydrocarbon class. Since 1977, the UAM has employed various versions of the Carbon-Bond Mechanism (CBM) developed at Systems Applications as part of a continuous kinetic mechanism research program sponsored largely by EPA.

Because of the association of reactions and reactivities with carbon bonds, the range of reactions and rate constants to be treated in a kinetic mechanism can be narrowed considerably when each carbon atom is treated according to its bond type. This concept was the basis for the first version of the Carbon-Bond Mechanism (CBM-I), a 32-reaction mechanism described by Whitten and Hogo (1977) and Killus et al. (1977). In this mechanism, the carbon atoms of each organic compound are assigned to one of the following groups: single-bonded carbon atoms, very reactive double-bonded carbon atoms, moderately reactive double-bonded carbon atoms, and carbonyl-bonded carbon atoms. The CBM-I scheme was employed in UAM applications in the late 1970's to Los Angeles, Sacramento, and Denver (Tesche and Burton, 1978; Reynolds et al., 1979). The CBM-II mechanism provided a better representation of the chemistry of aromatics and has been used in model applications to Los Angeles, Tulsa, Philadelphia, Kern County, the South Central Coast Air Basin, and San Luis Obispo during the first half of this decade (Tesche et al., 1981a,b; Killus et al., 1981; Reynolds et al., 1982; Tesche et al., 1983a,b; Haney and Braverman, 1985; CARB, 1985; Whitten et al., 1985; Haney et al., 1986; and Yocke et al., 1985). Recently, the CBM-III mechanism has been developed, which includes further improvements in the treatment of aromatics and updated reaction rate constants. CBM-III has been implemented in the UAM and is being used in the most recent applications to Los Angeles (Hogo and Yocke, 1987).

In 1976, six reactions describing SO_2 oxidation and four reactions accounting for the formation of sulfate, organic, and nitrate aerosol products were added to the UAM (Reynolds et al., 1976; Killus et al., 1977). Since that time, various investigators have incorporated improved treatments of sulfate and other aerosol formation processes (Tesche and Burton, 1978; Hogo and Yocke, 1987).

TREATMENT OF ADVECTIVE POLLUTANT TRANSPORT

Pollutants are transported primarily by advection, that is, by the mean or bulk motion of the wind. A major difficulty in treating advection is wind shear, the variation of the wind velocity with height above the ground. The wind speed near the surface typically increases with height, but the wind several hundred meters aloft may be faster or slower--and frequently from a different direction. In the original version of the UAM, wind shear was neglected, but later sensitivity studies indicated that inclusion of wind shear effects in the model could significantly alter NO_2 and ozone predictions (Reynolds et al., 1976). Consequently, the capabilities of the model were expanded to treat wind shear phenomena. Theoretical wind shear relationships were derived by Lamb (1976) using the predictions of a planetary boundary layer model developed by Deardorff (1972). These relationships were employed during the late 1970's when wind measurements aloft were not available. When wind measurements are available both at the surface and aloft, objective techniques are used to prepare appropriate three-dimensional wind inputs to the model.

Examples of the use of such techniques are described in the reports by Reynolds et al. (1979), Killus et al. (1981), Tesche et al. (1983a,b), and Haney et al. (1986).

TREATMENT OF TURBULENT DIFFUSION

Pollutants are transported and dispersed largely by the action of the wind rather than by molecular diffusion. The treatment of advection discussed above does not represent pollutant transport completely because it ignores the influence of small-scale features of the wind, called eddies. Describing these microscale features of the wind deterministically is difficult because available data are insufficient to characterize winds on such a fine scale. Consequently, it is necessary to parameterize the microscale turbulent diffusion processes. In the UAM, and most other grid models as well, transport of a pollutant by turbulent diffusion is assumed to be proportional to the rate of change of concentration with position (i.e., concentration gradient). The proportionality factor is the eddy diffusivity coefficient. These coefficients are sometimes treated as constants, but they can more accurately be treated as functions of atmospheric stability, surface roughness, and height above the ground.

Because of the empirical nature of the eddy diffusivity concept, it has heretofore been difficult to obtain precise measured or theoretical estimates for the diffusivity coefficients. A novel procedure for estimating these coefficients was developed by Lamb (1976) and Lamb et al. (1977) using control theory techniques in conjunction with the results of a sophisticated planetary boundary layer model to generate "optimal" diffusivity values. These algorithms for neutral and unstable stability conditions were used in UAM applications during the late 1970's. Subsequent modifications to enable improved treatments of stable conditions and the determination of the stability class from meteorological data typically collected in urban areas were added in the early 1980's.

TREATMENT OF OTHER MICROSCALE PHENOMENA

Several phenomena besides turbulent diffusion have characteristic spatial or temporal scales too small for explicit, deterministic resolution in urban-scale grid models. Two important examples are the rise of a hot, buoyant emissions plume from a large point source and the scavenging of ozone near roadways by NO emitted from vehicles. Studies of these phenomena and others have resulted in the incorporation of a state-of-the-art treatment of plume rise from point sources and a module to treat the microscale effects of the emissions from a network of line sources (e.g., highways) in the lowest layer of grid cells. In this layer, taken to be about 20 meters deep, NO emissions from vehicles on a roadway are allowed to react only with the ozone in the immediate vicinity of the roadway. More detailed descriptions of these algorithms and other microscale phenomena are given by Lamb (1976) and Lamb et al. (1977). Because these microscale phenomena are very difficult to treat analytically, only a preliminary version of the microscale module could be implemented and tested. This module has been used in only a few UAM applications.

In an effort to include a better representation of chemical reactions that occur in plumes at the subgrid scale immediately downwind of large point sources, Seigneur et al. (1983) embedded the Systems Applications Reactive Plume Model (RPM) in the UAM. Named the Plume-Airshed Reactive Interactive System (PARIS), this model treats the initial dispersion and reaction of emissions from large point sources using the RPM module. When the plume size becomes comparable to that of a grid cell, the pollutant mass within the plume is transferred to the grid model for subsequent simulation of transport and reaction. The PARIS model has been used in recent applications to Los Angeles, the South Central Coast Air Basin, and San Luis Obispo (Haney et al., 1986; Yocke et al., 1985)

TREATMENT OF SURFACE REMOVAL PROCESSES

Many types of pollutants, including NO_2, ozone, and SO_2 in particular, can be removed from the atmosphere by chemical reaction, adsorption, or absorption at various surfaces found in urban areas. A methodology for calculating the removal of gaseous pollutants by surface sinks has been incorporated in the model. Surface removal is assumed to take place in two steps: transport to the surface followed by uptake by the surface. Parameterization of this two-step process is accomplished by defining a resistance to mass transport and a resistance to surface removal. The transport resistance is estimated from theoretical considerations of turbulent transfer in the atmospheric boundary layer, and the surface resistance is obtained from experimental data on the uptake of pollutants by various types of surfaces (Killus et al., 1977).

OTHER TECHNICAL FEATURES

The numerical procedure used to solve the conservation-of-mass equations for each pollutant is an important component of a grid model. To minimize the propagation of numerical errors, several techniques have been studied for treating horizontal (advective) pollutant transport--an element of the model particularly difficult to handle numerically. The results of these studies indicated that, of the techniques examined, the SHASTA integration scheme provided the best balance between accuracy of prediction and computational speed (Reynolds et al., 1976; Killus et al., 1977).

The Urban Airshed Model contains features designed to reduce cost by minimizing labor and computing expenses. For example, computerized data preparation programs are available for automating the conversion of various types of data into the proper input format. These features have no effect on the model's treatment of atmospheric processes, but they do aid the user in preparing the inputs required by the computer programs.

The model output consists of the predicted pollutant concentrations for every grid cell. These predictions are generally averaged over a period of one hour and are saved on a computer file for subsequent display and analysis. As illustrated in Figure A-2, the predictions and actual measured concentrations can be plotted to aid in the evaluation of model performance. Figures A-3 and A-4 illustrate two possible means of displaying the spatial distribution of ground-level concentration predictions. Gridded census data can be used in conjunction with the model predictions to prepare population exposure estimates. Of course, a variety of other custom displays and analyses can be prepared depending on the needs of the user.

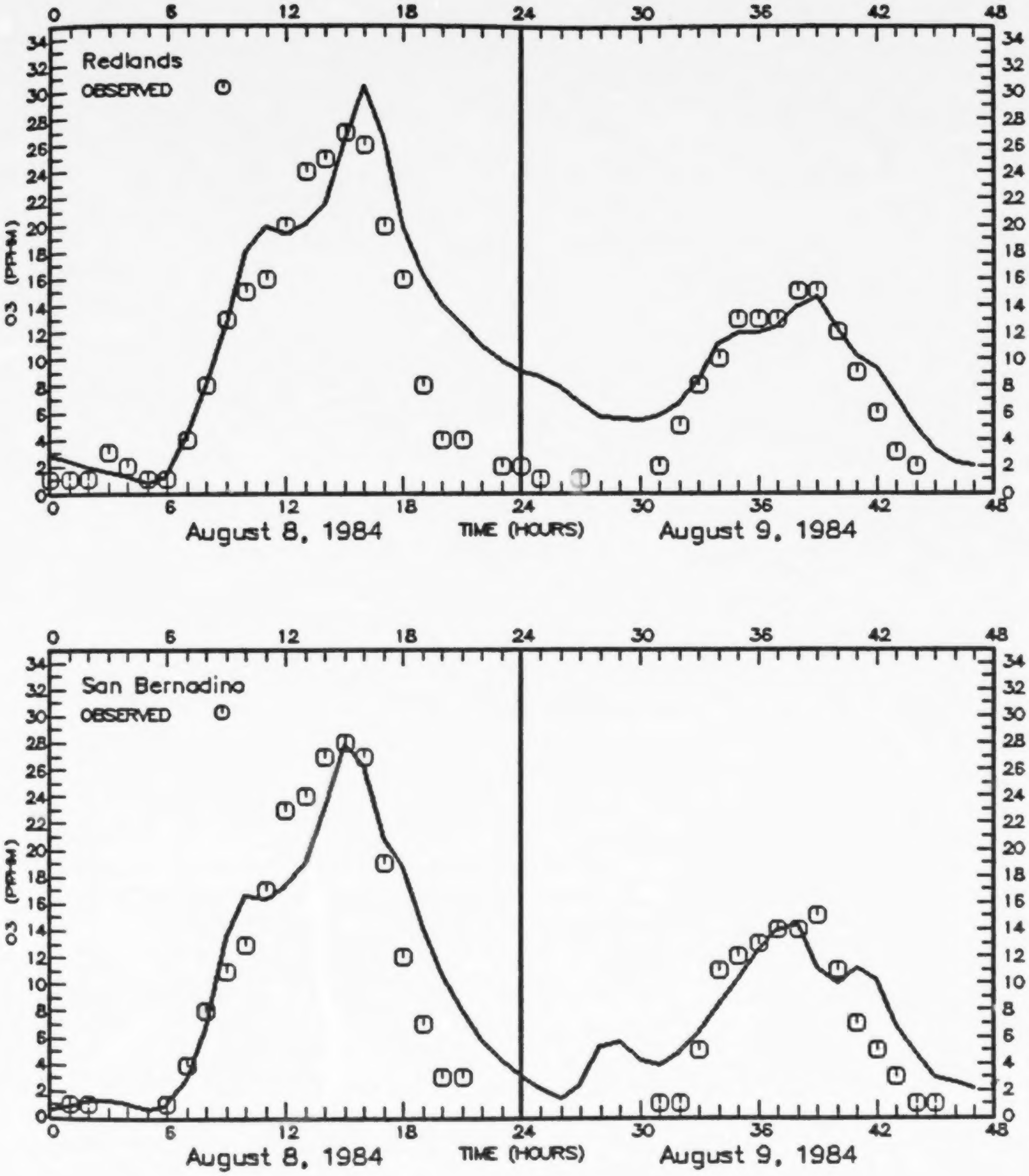
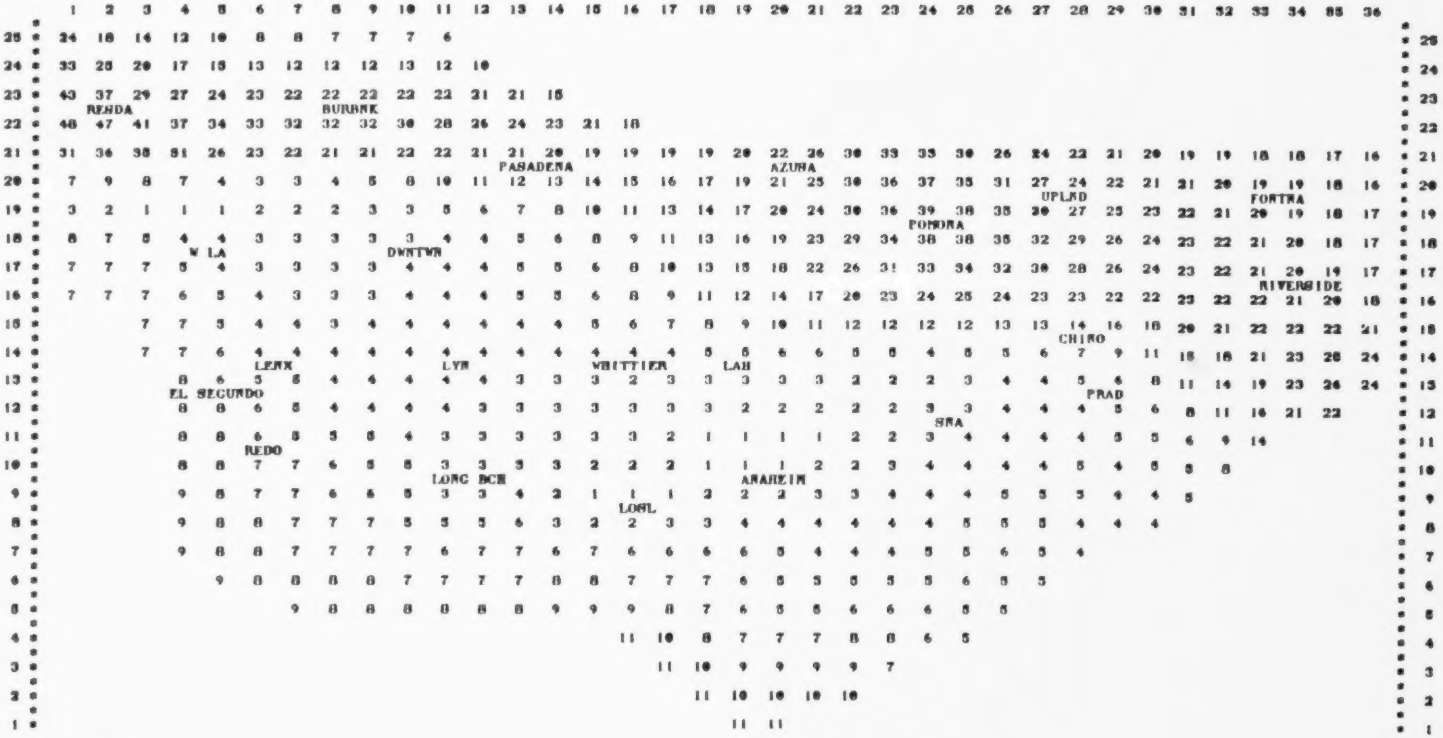


Figure A-2. Example of plots illustrating comparisons between computed and measured ozone concentrations (pphm) at air quality monitoring stations in the South Coast Air Basin for 8 and 9 August 1984. (Source: Hugo and Yocke, 1987.)


RENDA
BURBNK
PASADENA
AZUHA
UPLND
FONTNA
POMONA
W LA
DWNTWN
RIVERSIDE
CHINO
LENX
LYN
WHITTIER
LAH
EL SECUNDO
PRAD
SNA
REDO
LONG BCH
ANAHEIM
LOSL

Source: Tesche and Burton (1978).

Figure A-3. Example of a printed map of ground-level concentration predictions. Average ground-level O_3 concentrations (pphm) are given for Los Angeles on 4 August 1975 between 1:00 p.m. and 2:00 p.m. PST.

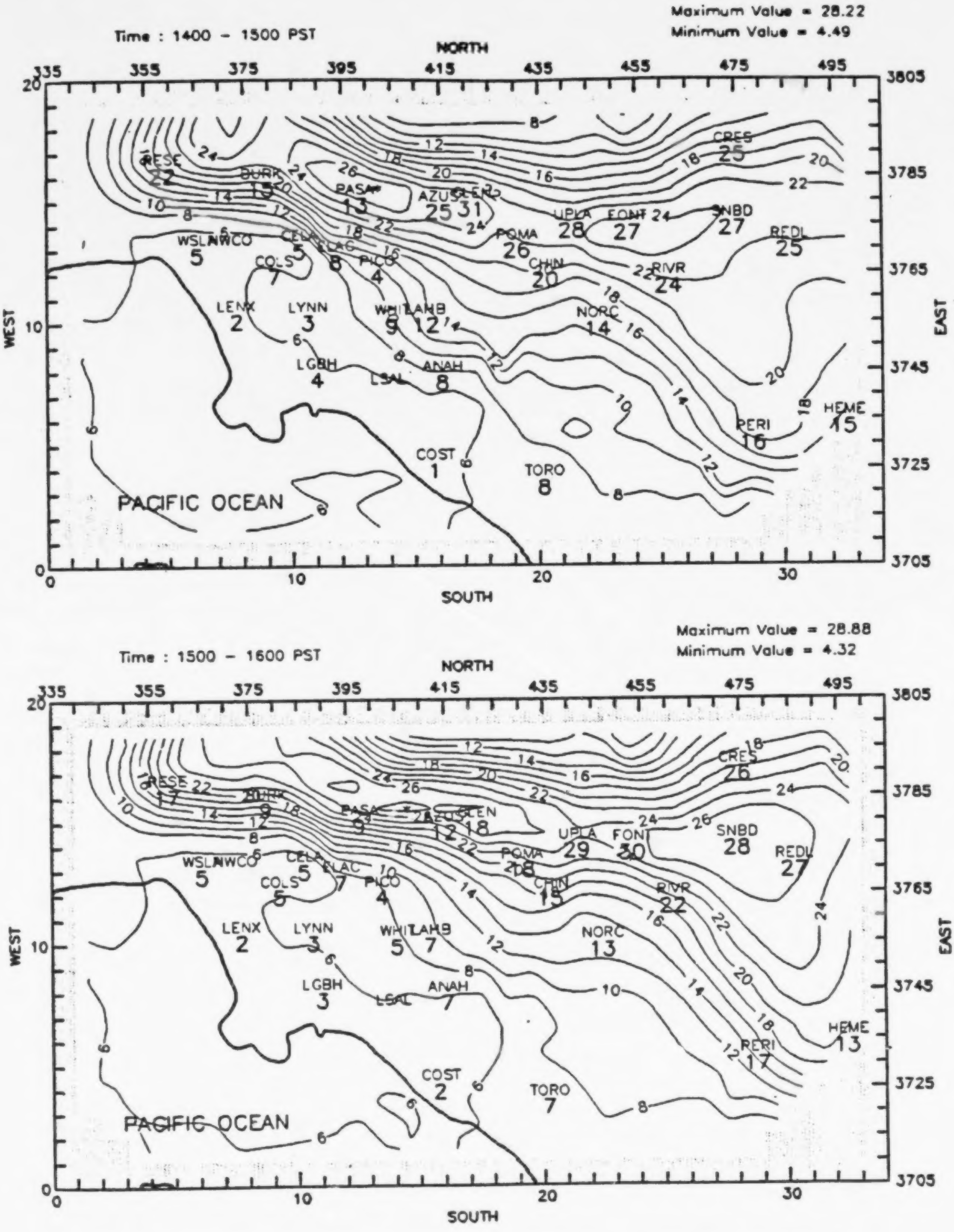

Time : 1400 - 1500 PST
Maximum Value = 28.22
Minimum Value = 4.49
NORTH
SOUTH
WEST
EAST
335
355
375
395
415
435
455
475
495
3805
3785
3765
3745
3725
3705
0
10
20
30
RESE 22
BURK 15
PASA 13
AZUS 25
GLEN 31
CRES 25
UPLA 28
FONT 27
SNBD 27
REDL 25
WSLA
NWCO 5
CELA 5
ELAC 8
PICO 4
COLS 7
POMA 26
CHIN 20
RIVR 24
LENX 2
LYNN 3
WHIT 9
LAHB 12
NORC 14
LGBH 4
LSAL
ANAH 8
PERI 16
HEME 15
COST 1
TORO 8
PACIFIC OCEAN
Time : 1500 - 1600 PST
Maximum Value = 28.88
Minimum Value = 4.32
NORTH
SOUTH
WEST
EAST
RESE 17
BURK 9
PASA 9
AZUS 12
GLEN 18
CRES 26
UPLA 29
FONT 30
SNBD 28
REDL 27
WSLA
NWCO 5
CELA 5
ELAC 7
PICO 4
COLS 5
POMA 18
CHIN 15
RIVR 22
LENX 2
LYNN 3
WHIT 5
LAHB 7
NORC 13
LGBH 3
LSAL
ANAH 7
PERI 17
HEME 13
COST 2
TORO 7
PACIFIC OCEAN

Figure A-4. Example of a plotted contour map of ground-level concentration predictions. Average ground-level O_3 concentrations (pphm) are given for Los Angeles on 8 August 1984 between 2:00 p.m. and 4:00 p.m. PST. (Source: Hogo and Yocke, 1987.)

APPENDIX B

STATISTICAL MEASURES OF MODEL PERFORMANCE

The primary objective of a model performance evaluation effort is to ascertain how well calculated air quality levels agree with those actually measured in the study area. For the Urban Airshed Model, this is usually accomplished by comparison of available hourly averaged measurements of ozone, NO_2, and other pollutants of interest with corresponding predictions for surface-level grid cells containing the monitoring sites. Several means of quantifying model performance have been developed, such as those reported by Bencala and Seinfeld (1979), Hayes (1979), Fox (1981), and Layland and Cole (1983). The measures most frequently employed in previous UAM performance evaluations have been those developed by Hayes (1979). In particular, Hayes identified five important attributes of desirable model performance:

1. accuracy of the calculated peak concentration,
2. absence of systematic bias,
3. absence of gross error,
4. temporal correlation, and
5. spatial alignment.

In Chapter III of this report, we discuss the results of several previous UAM performance evaluation studies. Particular attention is given to identifying historical trends in the overall bias and (gross) error in UAM results. This appendix provides a general discussion of the five performance attributes cited above an gives the formulas employed in the calculation of the bias and error figures presented in Chapter III.

ACCURACY OF THE CALCULATED PEAK CONCENTRATION

The accuracy of peak concentration predictions can be established by comparisons of the highest measured and predicted values at each air monitoring station. In addition, the highest measured concentration on a simulation day can be compared with: (1) the highest hourly averaged concentration calculated at the peak station, (2) the highest computed concentration at any monitoring location, or (3) the highest predicted concentration in any ground-level grid cell.

ABSENCE OF SYSTEMATIC BIAS

Absence of systematic bias refers to the desirable attribute of a model to neither underestimate nor overestimate pollutant concentrations. We calculated estimates of the mean bias for each day in the following manner:

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^{N} \frac{C_{m,i} - C_{c,i}}{C_{m,i}}$$

where: C_c and C_m are the computed and measured concentrations, respectively, and
N is the total number of comparisons.

The discrepancies between computed and measured values in the bias measure cited above are normalized by the measured value. Note that the bias can be calculated as a non-normalized quantity or, as in a few recent UAM studies, normalized by the average of the predicted and measured values. In addition, the median and standard deviation of the bias measure are also sometimes reported in performance evaluation studies.

Because of the uncertainties in measured values at low concentrations can be quite large (on a percentage basis), overall bias estimates are usually calculated only for those cases in which the observed and/or predicted concentration of ozone or NO_2 exceeds a value of from 0.02 to 0.05 ppm. Additional concentration cutoffs are sometimes selected to assess the possible bias at relatively high ozone concentrations. To further illustrate model bias characteristics, the normalized bias measure is plotted as a function of measured concentration. Such graphs provide a convenient means of displaying concentration ranges for which the model tends to overpredict or underpredict the measured values.

ABSENCE OF GROSS ERROR

The absence of gross error can be determined through the examination of the absolute value of the discrepancies between the predicted and measured values, either normalized by the measured value (or average of the measured and predicted values). The measure of error cited in Chapter III of this report were calculated using the following expression:

$$\text{Error} = \frac{1}{N} \sum_{i=1}^{N} \frac{| C_{m,i} - C_{c,i} |}{C_{m,i}}$$

If the error (or mean absolute normalized deviation) of the pairs of computed and observed values is small, then the model is said to exhibit "skill" as a predictor. If the mean absolute deviation is large, the model suffers from the presence of large gross errors. As with the bias determination, these measures are usually computed for ozone and NO_2 for all cases in which the measured values exceed an 0.02 to 0.05 ppm threshold. Additional computation are sometimes made for ozone for measured concentrations exceeding 0.12 ppm or, in Los Angeles, even higher cutoffs. Graphical representation of the error measure for ozone and NO_2 may be plotted as a function of measured concentration to complement the corresponding bias displays discussed previously.

TEMPORAL CORRELATION

Temporal correlation refers to the "timing" or "phasing" of the observed and computed ozone levels at a specified station. The temporal correlation for a given station can be determined by using the hourly pairs of observed and calculated concentrations to define the appropriate mean values. A correlation coefficient can then be calculated according to an expression given by Hoel (1962). Lack of temporal correlation can be ascribed to one or more causes, including inadequate characterization of emissions, wind, or mixing depth inputs. In addition, inaccuracies in the representation of atmospheric processes (such as chemical transformations) can also contribute to timing problems.

SPATIAL ALIGNMENT

The spatial alignment of observed and calculated concentration fields is another useful measure of model performance. For a given hour, imagine two concentration isopleths, one constructed from observed pollutant concentrations, and the other from the corresponding model calculations. If one isopleth were placed over the other, the degree of spatial misalignment would be easy to discern. Spatial alignment can be quantified by considering a sequence of "time slices." The calculated and measured concentrations for a particular hour are employed to calculate a correlation coefficient using a procedure described by Hoel (1962). Spatial correlation coefficients are computed for each hour, and estimates of the average coefficient for each simulation day are made following the procedure described by Hoel.

A value of the correlation coefficient approaching 1.0 signifies that the calculated spatial distribution of pollutants over the modeling region for a specific hour corresponds closely to that indicated by the measurements. Poor correlation is sometimes to be expected during hours when the measured values do not exhibit significant spatial variability. Although the absolute magnitudes of the predicted and measured values might be in reasonably close agreement, the small variations in pollutant levels from station to station might not be replicated in the variations of the predicted values, thus yielding a relatively low spatial correlation coefficient. Common sources of spatial misalignment include discrepancies between modeled and observed wind velocities, inaccuracies in the emissions inventory, and improper treatment of photochemistry.

Graphical representations of results are also a useful tool in evaluating the model's ability to replicate spatial concentration patterns. Visual examination of plots showing both model predictions and actual observations provides insight into possible model performance problems. Similarly, spatial maps of the computed ozone ground-level concentration field illustrate potential spatial alignment difficulties.

APPENDIX C

SURVEY ON PHOTOCHEMICAL AIR QUALITY MODELING

Part 1 of this appendix contains the questionnaire distributed to each interviewee prior to the telephone interview.

Part 2 of this appendix provides a list of the interviewees.

PART 1

SURVEY ON PHOTOCHEMICAL AIR QUALITY MODELING

SURVEY OBJECTIVE

The Regional Acid Deposition Model (RADM), whose development is being sponsored by the National Acid Precipitation Assessment Program (NAPAP), will eventually be used to estimate the effects of control strategies on acidic deposition. The model's performance will be assessed by comparison of model estimates to measurements obtained from a data base designed for model evaluation. Presently, NAPAP is attempting to assess the importance of the quality of model performance to policy makers. Our approach is to examine an historical precedent. The present study, of which this survey is a part, is intended to analyze the acceptability of a grid-based photochemical air quality model, and the model estimates, to those in the policy community of California over a ten-year period. Your experience with air quality modeling will help us identify key policy-related modeling issues and thus help guide the RADM evaluation process.

This survey focuses on the use by policy makers of the estimates derived from the Urban Airshed Model (UAM), developed by Systems Applications Inc. (SAI) for simulating ozone formation on an urban scale. We are specifically interested in your views on the contributions of model estimates to decisions made during the course of four policy-making activities in the Los Angeles area. Any additional comments pertaining to applications of the model outside the South Coast Air Basin would also be welcome. If, for any of the activities, you feel there was a difference between your personal views and views that you think were representative of the institution with which you were affiliated at the time of the activity, we would appreciate your best assessment of the institutional view.

The survey will be conducted by telephone. A copy is enclosed here for your convenience. If you would prefer to return written responses, you may so inform us when we contact you by telephone.

SURVEY QUESTIONS

1) Nature and Extent of Involvement

The first set of questions pertains to the nature and extent of your involvement in four specific activities in the Los Angeles area. These activities are Rule 1135 (1979), two Air Quality Management Plans (1979 and 1982), and the Berg subcommittee hearings on the need for additional NOx controls (1985).

a) How would you describe the nature of your participation in Rule 1135 (more than one choice may apply)?

- You did not participate.
- You attended hearings or had some interest in this activity but did not directly participate.
- You prepared or helped prepare a report, testimony, or other material presented at hearings.
- You testified at hearings.
- You participated in a decision made at hearings.
- Other (please explain).

b) Using the preceding list, how would you describe the nature of your participation in the 1979 AQMP?

c) Using the list above, how would you describe the nature of your participation in the 1982 AQMP?

d) Using the list above, how would you describe the nature of your participation in the 1985 Berg subcommittee hearings?

2) Familiarity with the UAM

For each of the activities described above, air quality estimates derived from the Urban Airshed Model (UAM), developed by Systems Applications Inc. (SAI), were reported. The following set of questions pertains to your familiarity with the UAM and the air quality estimates derived from the UAM.

a) At the time of Rule 1135, how would you characterize your level of familiarity with air quality estimates derived from the UAM?

- You were unaware of the UAM estimates.
- You were unfamiliar with the UAM estimates but aware of their existence.
- You heard the UAM estimates described at a hearing or in some other presentation.
- You read a written summary of UAM estimates.
- You read one or more reports in which UAM estimates were described in detail.
- You prepared a summary of the UAM estimates.
- Other (please explain).

b) Using the preceding list, how would you characterize your level of familiarity with air quality estimates derived from the UAM at the time of the 1979 AQMP?

c) Using the list above, how would you characterize your level of familiarity with air quality estimates derived from the UAM at the time of the 1982 AQMP?

d) Using the list above, how would you characterize your level of familiarity with air quality estimates derived from the UAM at the time of the 1985 Berg subcommittee hearings?

3) The Importance of Performance Evaluations

The next set of questions focuses on the impact that performance evaluations of the UAM have had on your thinking. By performance evaluation, we mean comparison of model estimates to a set of measured ozone concentrations at a number of monitoring sites on one or more days.

a) Had you seen any performance evaluations of the UAM at the time of:

- Rule 1135?
- AQMP 1979?
- AQMP 1982?
- 1985 Berg subcommittee hearings?

(If your answer for each activity is 'no', please skip questions '(b)' through '(e)' and continue with '(f)'.)

b) The measures of performance used in the performance evaluations have included:

- Average difference between estimated and measured ozone concentrations.

- Bias (systematic over- or underprediction).

- Ratio of peak estimated to peak measured ozone.

- Prediction of temporal and spatial trends.

- Percentage of hours with estimates within 50 or 25 percent of the measured values.

Which, if any, of these measures affected your assessment of the UAM performance? Which measure most affected your assessment?

c) At the time of the activity when you first saw a performance evaluation, how would you characterize the UAM performance levels?

- Performance was too poor for estimates to be considered seriously.

- Performance was poor enough that model estimates were questionable.

- Performance was good enough for model estimates to be considered.

- Performance was good enough to provide a high degree of confidence in the estimates of the model.

Did this performance level change for subsequent activities?

To which level? When?

d) Did any measures of performance other than those listed above affect your assessment of the UAM (please explain)?

e) Would you have liked to see other measures of performance included in the evaluations (please explain)?

f) Did the existence of performance evaluations, apart from specific details or aspects of the evaluations, affect your assessments of the UAM?

4) The Importance of Sensitivity Analyses and Other Measures of Reliability

The next set of questions pertains to factors other than performance evaluations that may have affected your assessment of the accuracy and reliability of UAM estimates.

a) For the first activity in which you were involved, which, if any, of the following affected your assessment of the accuracy and reliability of UAM estimates?

- Sensitivity analyses that showed what ozone levels the model predicted for hypothetical emission-reduction scenarios, such as the case of a 50 percent reduction in ROG and NOx emissions.
- Sensitivity analyses that demonstrated the degree of dependence of model estimates on initial or boundary concentration assumptions, or on possible uncertainties in wind or emissions input data.
- The consistency of model estimates with other information pertaining to the probable consequences of emission changes on air quality.
- Other sensitivity findings (please explain).

b) Which of the preceding factors, if any, affected your assessment of the UAM during subsequent activities?

c) Could you name any other influences on your assessment of model accuracy and reliability?

5) Impact of Model Estimates

The next set of questions pertains to the impact that the UAM estimates had on your thinking.

a) For the first activity in which you were involved, to what extent were your opinions affected by the UAM estimates?

- Little or not at all.

- Marginally to moderately.

- Substantially: UAM estimates reinforced, altered, or otherwise influenced your opinion on one or more issues, although other influences may have been as great or greater.

- More substantially than by any other factors.

b) For each subsequent activity in which you were involved, how would you characterize the influence of the UAM using the preceding list?

c) For the first activity in which you were involved, which, if any, of the following would have changed your answer to question '(a)' above?

- A greater degree of accuracy and reliability, as shown by the performance evaluations.

- A greater degree of consistency between model estimates and other information available to you concerning the probable consequences of emission changes.

- A lesser degree of dependence of the model's estimates on the reliability of emissions data, wind data, or assumptions concerning initial or boundary concentrations.

- A greater number of evaluations, for varied meteorological conditions.

d) For each subsequent activity in which you were involved, which of the preceding would have changed your answers to question '(b)' above?

e) Did either the UAM or the UAM performance evaluations influence your thinking on any of the following issues?

- The effects of NOx emission controls on ozone concentration: an urban-scale nonlinearity issue.
- The length of simulations required to properly study the effects of emission reductions: the initial condition issue.
- The need for a data base specifically designed for model performance evaluation.
- The utility of grid-based photochemical models for addressing contributions of single sources to ozone formation.

6) The Influence and Utility of Photochemical Air Quality Models

The last set of questions pertains to your personal assessment of the influence and utility of photochemical air quality models to air quality decisions.

a) For the first activity in which you were involved, what do you believe was the extent to which the UAM estimates affected the thinking of others?

- Little or not at all.
- Marginally to moderately.
- Substantially.
- More substantially than any other influence.
- Cannot answer.

b) For each subsequent activity in which you were involved, what do you believe was the extent to which the UAM estimates affected the thinking of others (please use the preceding list)?

c) For each activity in which you were involved, could you name any factors that you believe were more important than the UAM estimates in influencing the thinking of others?

d) Photochemical air quality models can be used in a number of ways to aid in formulating decisions on air quality and emissions control. Some of these uses require a greater degree of confidence in the model than do other uses. If you were involved in modeling or air quality decisions in 1975, how would you characterize the uses that you felt were appropriate for photochemical air quality models at that time?

- Identifying information needs.
- Identifying air quality trends but not precise concentrations.
- Identifying possible outcomes of control measures that would require further investigation.
- Complementing other evidence on the outcomes of control measures.
- Assessing the outcomes of control measures for policy decisions.
- Other (please explain).

e) Do you yourself feel that there was any significant change in the appropriate uses of models between 1975 and 1980? If yes, how?

f) Do you yourself feel that there has been any significant change in the appropriate uses of models since 1980? If yes, how?

7) Additional comments

Thank you for your responses. We would welcome any additional comments that you might have to offer.

PART 2

LIST OF INTERVIEWEES

Dr. Praveen Amar, Manager
Acid Deposition and Aerosol Research Section
California Air Resources Board

Ms. JoAnne Aplet, Director
Planning Division
South Coast Air Quality Management District

Mr. Thomas Austin
Sierra Research

Mr. Bruce Beyaert
Chevron USA

Professor Glenn Cass
Chemical Engineering Department
California Institute of Technology

Professor Laurence Caretto
Mechanical Engineering Department
California State University

Mr. R. David DiJulio
Law Offices of R. David DiJulio

Mr. Kevin Golden
U.S. Environmental Protection Agency
Region IX Office

Dr. Alvin Gordon
University of California, San Diego

Mr. John Grisinger
Planning Division
South Coast Air Quality Management District

Mr. Andrew Gunther
Energy and Resources Group
University of California, Berkeley

Dr. Thomas Heinsheimer
Titan Systems

Mr. Michael Hertel
Southern California Edison Company

Dr. John Holmes, Chief
Research Division
California Air Resources Board

Mr. David Howekamp
U.S. Environmental Protection Agency
Region IX Office

Mr. Charles Kay
Texaco, Inc.

Mr. David Layland
U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards

Dr. Chung Liu
Planning Division
South Coast Air Quality Management District

Mr. Nader Mansour
Southern California Edison Company

Mr. Terry McGuire, Chief
Technical Support Division
California Air Resources Board

Mr. Vincent Mirabella
Southern California Edison Company

Ms. Mary Nichols

Mr. John Pearson
U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards

Mr. Andrew Ranzieri
Technical Support Division
California Air Resources Board

Mr. Gary Rubenstein
Sierra Research

Professor Robert Sawyer
Department of Mechanical Engineering
University of California, Berkeley

Ms. Sabrina Schiller
Coalition for Clean Air

Professor John Seinfeld
Chemical Engineering Department
California Instiute of Technology

Dr. Bernard Steigerwald, Director
Regional Programs Office
U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards

Mr. J. E. B. Stuart

Mr. William Sylte, Chief Deputy Executive
Officer
Executive Office
California Air Resources Board

Mr. Joseph Tikvart, Chief
Source Receptor Analysis Branch
U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards

Mr. Peter Venturini, Chief
Stationary Source Division
California Air Resources Board

Professor Arthur Weiner
University of California, Riverside

Dr. Stephen Ziman
Chevron USA

APPENDIX D

SUMMARY OF INTERVIEWS

This appendix presents detailed comments of interviewees.

In the master version of this appendix, each comment is preceded by a number, uniquely coded to an interviewee. In accordance with agreements concerning preserving anonymity, this coding has been removed. However, the master version has been preserved.

PERFORMANCE MEASURES UTILIZED BY INTERVIEWEES

Drive to look at model more deeply stemmed from realizing its complexity. In 1982, asked: Did it predict temporally and spatially? How far off was it? Did predictions fall within error and bias of 25 - 50 percent? Did predictions follow trends?

In 1979: predicted peaks in different portions of the basin.

From regulatory point of view, the principal measure utilized should have been ratio of peak estimated to peak measured, but was actually prediction of temporal trends. Never validated model for updated emission inventories and air quality.

Average difference between estimated and measured ozone (about 35 percent).

Ratio of peak estimated to peak measured ozone was very important: primarily interested in this measure because model was intended for regulatory policy.

Gross error and peak estimated to peak measured were most important. Bias and prediction of spatial/temporal trends were also important because effects were different in different parts of the basin. Some locations have more severe problem than others and you want to focus on problem areas.

Largely took results at face value.

Spatial and daily variations were approximately correct, which was reassuring. Concerned about ozone as a surrogate for total oxidant. No PAN, nitrates. Didn't predict NO_2 well. Downwind effects: far eastern basin and desert outside model boundary in time and space.

Used as many graphical representations as possible.

Clean Air Act standards implied need to look at maximum hourly concentrations in vicinity of monitoring stations.

Gross error for NO_x.

Most important measures from scientific perspective were gross error and bias; from regulatory perspective, peak predicted to peak measured. Not much emphasis on temporal trends. Percentage of hours with estimates within 50 or 25 percent of measured correlates with gross error and bias, and is easy for decision makers to understand.

All measures had impact - each tells something.

OTHER MEASURES OF PERFORMANCE CITED BY INTERVIEWEES

Scientific background was more influential than any specific measures.

Influenced by concerns that were never dealt with in the policy discussions. First, effect of boundary conditions. Second, assumptions involved in meteorology. Third, uncertainties in emission inventories. Fourth, locations of point sources in grid (which were not in centers as assumed).

Need to look at internal hydrocarbon-to-NO_x ratio and compare to actual.

Sensitivities.

General familiarity with modeling and informal discussions with modelers.

Sensitivity: performance evaluations were tied in with sensitivity runs. Without that, you would only have a feeling that model was close enough; you wouldn't know how much the answers change with respect to the inputs.

Hydrocarbon-to-NO_x ratio.

NO_2, which wasn't as detailed as ozone and was far off for concentrations and timing. Concerned that model may be OK for ozone at the expense of other pollutants.

I had heard general misgivings expressed by experts I trusted.

Comparison of specific application to applications in other places: relative scale of quality of application.

MEASURES THAT WERE DESIRED BUT NOT AVAILABLE TO INTERVIEWEES

Person-days above the standard and "person-day-ppm" (i.e., population exposure).

On grid model, points are compared to grid averages. Would like point measures compared to adjacent grid cells.

Would have liked to see more detailed discussion of effects of boundary conditions, meteorological assumptions, emission-inventory uncertainties, and locating point sources in centers of grid cells.

Would have liked to see comparison of mobile sources to stationary sources, which is the big issue.

Didn't need to see additional performance measures because had no confidence in the model anyway.

Evaluations were applied only to ozone. Performance needed for other key species: NO_2, organic species (aldehydes and oxygenated species were never checked).

Evaluations were needed for NO_2, hydrocarbons, and hydrocarbon-NO_x ratio.

Models can only be a guide, not a dictator.

Would like to communicate to decision makers that error band of 25 percent is not too big; compare this uncertainty to uncertainties of information that decision makers are now using.

INTERVIEWEES' CHARACTERIZATIONS OF UAM PERFORMANCE

In 1982, assumed you could get close to observations. In 1986, believe directions but not the numbers themselves. Need to look not only at peak but also the area under the curve. Must look at other pollutants: there is good reason to understand what the hydrocarbon:NO_x ratio is since if it's 4:1 or 8:1, hydrocarbon control is important. Don't have good measures of performance for hydrocarbons and NO_x, which are each influenced by the sources that emit them. Errors on their measurement may be larger than on their predictions.

Inadequate for quantitative decisions. In 1979, the model did not adequately predict NO_x. By 1986, the modelers quit reporting NO_x performance measures. The quality of the model improved substantially between 1979 and 1986, but it still is not adequate for quantitative assessment. The model is good for qualitative assessments, however. A fundamental problem in evaluating performance is that you can "fudge" to match observations on an existing day; therefore, the model's performance on a future day is much more uncertain.

Ozone was within 20 - 30 percent of measured values, with the time of occurrence within about one hour. NO_2 was not as good, and CO was not good. With incorporation of nighttime chemistry, level of confidence in the model increased to fourth level.

For all activities, the model was good for qualitative assessments. However, boards thought of the model as a black-box, quantitative tool. A problem was that the board expected more precision.

Performance was between the third and fourth levels. Can't look at absolute numbers. Value of models is that they provide a means for prioritizing control measures. Part of overall planning process is to constantly update models, evaluate control strategies, and compare to what's already been done. Given the uncertainties in emissions and monitoring, the model is surprisingly good.

By 1982 AQMP, had a better feeling for the model, based largely on sensitivity analyses. Up to third performance level but not yet the fourth.

Troubling that presentations of the model results did not discuss whether 35 percent gross error affected the conclusions.

Performance did not reach the fourth level because of spatial and temporal trends.

Initially skeptical of model in 1982. Assessment of performance went to fourth level after seeing performance evaluations.

In 1979, UAM was the best tool available, and was good enough to use. The alternative of EKMA is simplistic and unsatisfactory.

The question is: what is good performance? You may get good prediction in the near field or for the far field but not everywhere. Performance evaluations are important, but you need criteria for judging them. If goals are specified, you then can judge the performance evaluations.

Under the impression that model improved from 1979 to 1985, but not enough to increase performance to the fourth level.

Performance was good enough for model estimates to be considered if used in the right context: to test sensitivity of ozone to various strategies, and to examine relative effects.

Performance is between the third and fourth levels, and is good enough for control-strategy evaluation. About +/- 30 percent for peak ozone, leaning toward underprediction. Predicts time well but location differs from measured.

For JIMS, data base was poor. Underestimated grossly in isolated inland valleys. Although performance poor, could get qualitative insights. Couldn't use for SIP.

Neither ozone nor NO_x were well predicted. Also sensitivity - so sensitive that you had no confidence as a regulator. Confidence diminished by knowledge that many input variables had to be estimated. In neither 1979 nor 1982 did regulatory board feel comfortable with predictions. Never impressed by 3D models enough to rely on them for regulatory purposes. Most useful model was EKMA, which was more immediate and easier to work with for ozone. Thought to be as good as more sophisticated models. Personal position: should be supporting model development to get tool that is useful for detailed assessment of land use, automobile patterns, plant locations, etc. Boards' feelings based on performance evaluations for ozone and NO_x. Also based on sensitivity analyses: if you get much more than linear response, the model is too sensitive to give you much confidence in the results. Tempered by knowledge that models have improved. Seinfeld model now predicts nitric acid. More interest in speciation now. Goal is to know results within 10 - 15 percent with confidence.

Performance improved but always remained within the third level.

Model off 20 - 40 percent. Therefore, numbers were not very credible but sensitivity analyses OK. Use model qualitatively.

Had no basis for looking at model in 1979 and 1982. Really started looking at performance in 1984, when all models were evaluated and we could begin to judge measures quantitatively. In 1984, examined ozone, NO_x, hydrocarbons, and hydrocarbon:NO_x ratio. Ozone very good, NO_2 somewhat poor. Model underestimated hydrocarbon:NO_x at three stations by 60 - 80 percent; still evaluating this problem. Criteria: predicted/observed within 30 percent, gross error about 40 percent, normalized bias about 30 percent. Would like to see them tightened up. Good data might reduce errors. ARB management will argue that model is good for quantitative assessment. Looked at model to see if ozone standards would be achieved. Answer was "no". Then looked at alternative strategies.

Performance improved from 1979 to 1985. Not because model was better but because more inputs were used - reinjected ozone aloft. Still overpredicted in coastal zones and seriously underpredicted in eastern portion. Doesn't do a wonderful job predicting peaks. Ratio of peak predicted to peak measured is, in a certain sense, an outlier. Models don't do well with outliers. But, in the regulatory framework, you need to look for peaks. On average statistics, UAM isn't too bad. Performance was somehow only mentioned in passing; too technical for board.

It's my understanding that, because of certain failings (2nd and 3rd day "slosh", nighttime chemistry, etc.), the results could not be considered totally reliable.

Confidence in model and ability to get data have improved since 1979. Hope to see model reach the fourth performance level by the end of 1987. Model validates reasonably well compared to others.

Model performance depends on what you want to do with the model. Generally, model has performed too poorly for quantitative assessment: predicting future ozone is beyond the capabilities of the model. Can get crude estimates of sensitivity to large-scale hydrocarbon or NO_x changes, and qualitative assessment of different scenarios. Performance varies with application. Even in Kern County (one of the better applications), model performance was better for some purposes than others. In Kern, the conclusion of high confidence was that ozone would respond to NO_x as well as hydrocarbon control.

INTERVIEWEES' ASSESSMENTS OF THE NEED FOR EVALUATIONS, APART FROM SPECIFICS

Depends on how the evaluation is done. Does everyone agree on whether evaluation is proper? Is the emphasis on matching hydrocarbon-to-NO_x ratio important?

Need to go into the details for credibility.

Without performance evaluations, couldn't accept predictions.

Fact that some evaluations were done suggested validity; partial presentation was bothersome.

Added to general credibility.

Credibility is number-one concern.

Had to be done.

More credible as a result of performance evaluations.

Absolutely essential.

Would not have been credible without.

Absolutely necessary.

Absolutely necessary.

Good to have them.

Model of no value without performance evaluations.

Absolutely not credible without.

Could not be credible without. But contrast Gaussian, where you never do performance evaluations and the results are always accepted.

Cannot avoid performance evaluations - otherwise the model is a black box.

Absolutely.

No, because already knew how to balance off the evaluations against what we already knew.

COMMENTS CONCERNING SENSITIVITY ANALYSES

Were helpful. Relied on them, particularly for Kern county. Helpful for forming qualitative conclusions.

Initial conditions biggest problem. Had surface data only - none aloft. Wind-field calculations depend on number of stations and their placement in the airshed. Need more wind data, or better method than objective interpolation scheme.

For 1979 AQMP, sensitivity analyses showing effects of 50 percent reductions were important because primary purpose was to develop a plan. The consistency of model estimates with corroborating evidence, such as trends, is also important.

Useful for giving insight into way the atmosphere behaves, but couldn't recognize a sensitivity result that shouldn't occur. Need distinction between sensitivity runs and control strategies. Real control measures are not as simple as the sensitivity analyses that have been done. Need a way to generate realistic scenarios for future emissions, which isn't easy since you need to know growth scenarios. Now, we discredit anything that doesn't look like a real control measure. Sensitivity analyses should be part of validation.

Very important tool. Given uncertainties, sensitivity analyses say whether or not a factor is important. Have been concerned about lack of comparison of model results to present. Runs done for 1974 but were not calibrated to model the changes in emissions. Could argue that application to other geographical areas is a test, but should try 1984.

Only sensitivity analysis presented was the one concerning reductions in emissions, and that was presented as a result (not by SAI). Led to looking for other sensitivity analyses that weren't presented. Should allocate money for sensitivity analyses. Need to put all sensitivity analyses together, and do sensitivity analyses before going public. Need to do all this before results gotten because people "get on board" depending on the results.

Sensitivity analyses really convincing - as important as performance evaluations. Good to see similarity to gross data: stations not responding to NO_x controls were in areas close to freeways.

Fifty percent reduction wasn't meaningful in terms of specific control strategy. From perspective of SCE, 50 percent decrease didn't mean anything because didn't know if interests affected. From perspective of someone who wants improved air quality, didn't know if realistic. Fifty percent reduction is somewhat academic: implication is lost until you identify the target.

Biggest utility of model: can change ratios of reductions in various pollutants. Gives direction of change pretty comfortably. Guides broad strategy.

District has assumed all emission reductions equally good. ~~Good~~ model could tell you which reductions most effective. Model could be made more effective. Used scenario of 50 percent reduction because had been told that it would take that much to show an effect. Hoped model would show smaller changes: extrapolation down to smaller numbers would have been better (10 percent ideal).

Reassuring that sensitivity analysis of 50 percent emission reductions showed that things went the direction that you expected. Affected thinking about appropriate regulatory policy. Initial thought: control both hydrocarbons and NO_x equally aggressively. Model softened feeling.

Looking at sensitivity analyses more important than performance evaluations. Sensitivity analyses are a part of performance evaluation: if you take away emissions and pollutant concentrations don't collapse, model is in trouble. A lot of models may reproduce past history. Models are deterministic, and much input data may not be available. Interpolation schemes are therefore needed (need to create wind fields aloft; also, no measurements for pollutant concentrations above the boundary layer). Many important inputs therefore estimated. Running model gives latitude: can set up model so that it does a credible job on one given day. Need to run model for several different days with same set-up procedure. If you had warped inputs to fit a given day, sensitivity analyses would have shown it.

Sensitivity analyses not necessarily consistent; didn't do same analyses everywhere. Looked at hydrocarbon control across studies. NO_x control not significant concern. Where is NO_x control effective? Low hydrocarbon:NO_x ratio, then NO_x control not effective. Issue is in doubt for high hydrocarbon:NO_x ratios.

Range of model performance important. Would suggest tests of winds to accurately reflect terrain, sea-breeze patterns. Other sensitivity findings include mixing height and number of vertical cells (especially if there are elevated point sources).

In 1979 and 1982, rollback model and EKMA just as good and easier to use. Uncertainties in emissions (inventory problem) in fairly poor shape. Still a long ways from having problem under control. Also overly sensitive to boundary conditions.

Sensitivity analysis of fifty percent emissions reductions very useful; to District, they were most important. Can't predict absolute values but useful for relative changes. Therefore, across-the-board fifty percent reduction helpful in assessing impact. Small reduction change not helpful. Get feeling for how much change you get from what emission changes. Troubling part was sensitivity to initial and boundary conditions. Highly dependent on initial and boundary conditions in Los Angeles. Initial and boundary conditions always a concern since 1979. Consistency has proved itself over years, yielding more confidence. Emissions pretty robust. Three-dimensional structure of winds is sensitive, and no data available. Less a sensitivity problem and more a need for more accurate data. Not much feeling for how sensitive model is to wind.

JIMS - day simulated ozone was 7 - 8 pphm. Impact of increment was to increase to 8 - 9 pphm. Need simulations for above-standard days. Kern - most enlightening analyses were to shut off west side: didn't affect the east side, which implies two air basins. This was with and without NO_x controls. Northern boundary concentrations high, which implies transport real and you get less than proportional changes in air quality from emissions reductions within domain.

In 1981, various combinations of hydrocarbon and NO_x reductions were important. Now, for diagnostics, need sensitivities: meteorology (winds), boundary and initial conditions, emission uncertainties. May be underestimation of hydrocarbon inventories; should maybe be increased two to four times. However, as modeler, you have total freedom to adjust anything except emission inventory, which may be the most uncertain. In Kern, what is the effect of upwind transport into county, upslope flows, etc.?

In 1979, sensitivity analyses not extensive or not greatly advertised; maybe buried in report. So expensive to run in 1979, runs limited. In 1982, very sensitive to initial and boundary conditions. Take care of initial conditions by running model a long period of time. Gives more confidence in results. Multi-day runs came later on. Also looked at emissions; counter to experience or counter-intuitive.

I was particularly affected by the ARB's presentation of a cause-effect relationship between NO_x reductions at the coastline and ozone reductions inland. This info was admittedly statistical, but it seemed to indicate concentrating on hydrocarbon reductions for best results may not be the only approach. The statistical results, as addressed by the ARB, appeared to contradict the SAI predictions. This, just coupled with the logic of the need to predict under various conditions, under conditions which consider the real world, cast doubts in my mind as to the extent to which we could depend on the SAI model.

The two-day 1974 air pollution episode (classic worst case) used in model; often attacked (need 3 - 4 days). Affected judgement on what needs to be done to improve models.

In 1979 - 1982, would have taken greater than 75 percent control to attain standards - didn't need AIRSHED to tell us this. Historical trends more convincing. Second most important: if model blows early morning hydrocarbon and NO_x, other answers very sensitive.

OTHER INFLUENCES ON ASSESSMENT OF MODEL ACCURACY AND RELIABILITY

Very sensitive to boundary conditions. Is this an artifact of the model? May be real in Los Angeles. Model is dealing with pphm. A source up to one percent of total would involve ppb. At limits of measurement technique. May be spending money to reinvent calculus: very small changes can't be detected. Cutting emissions by one, ten, or twenty percent only changes concentrations by pphm.

Non-chemical sensitivities: plume rise, wind speed, etc. Also politics. Decision makers who rely on model may be scientifically unknowledgeable and uncomfortable with it. They need to make decisions for control strategies that will actually be beneficial: need to reduce ozone relative to net population exposure, need to not underestimate controls. Depending on politics, lean one way or the other. SOCAB tend to be conservative: weren't going to buy not implementing NO_x controls. Santa Barbara against oil development. Kern county concerned with cost of controls: may not get anything out of it.

Little aspects so esoteric that only the modelers talk about them strongly affect the answers. Include boundary conditions and initial assumptions. Publicly characterized as "niggling differences", the "little" assumptions had overwhelming impact on conclusions. For example, the hydrocarbon-to-NO_x ratio of 30 or 40 to 1 dictated the outcome that hydrocarbon controls played no role and only mitigation controlling NO_x at project site would mitigate ozone.

Indirectly, how were regulatory districts using the model. If they weren't receptive, couldn't use the results. Now, it's the tool to be used.

Have only one meteorological regime (June 25, 1974). Single most limiting factor. Have since had population shifts, etc. Haven't collected data (people will pay to run but not improve the model). Want to validate against 1984, 1987 baseline. Can put in new inventory. Negative image from everyone saying only one-day run. Need representative bad days.

None.

Political influences. Partisanship made model results suspicious. Need to keep sensitivity work as ivory-towered as possible. Didn't like not seeing the other sensitivity analyses.

Concerned that developer of model appeared to testify on behalf of industry clients, though he didn't say anything he wasn't backing up as a scientist. Although no evidence of influence, this made me uncomfortable and left me less confident in the model. Limiting discussion to ozone may be limiting the usefulness of the model as a tool. Specifically, NO_2, nitrates, visibility, and acid deposition. Question: did the lack of prediction of NO_2 affect judgement of model? Answer: yes, less willing to spend time to use model if just good for ozone.

Technical experts presented support.

No.

One problem: fact of simulating one day's occurrence. More recently modeling multi-day. Second day provides boundary conditions and input.

Degree of treatment of physics and chemistry.

Intangibles: quality and scientific honesty of people developing the model were an important consideration. SAI went out of their way to maintain integrity, take a balanced view, and not take extreme positions when surrounded by people who were extreme. Actual reluctance on the part of the regulatory agencies because results might turn out contrary to previously held beliefs and policies. Maybe because they thought models weren't good enough and would be used against them. Attitude changed over time. At beginning, "us" versus "them". As model developed and took on more credibility, regulatory agencies decided they better get on board. During the 1975 to 1977 time period, SAI convinced community they were objective.

An assessment of theoretical soundness of scheme. Important to take a look at the underlying calculations. Judge whether model likely to blow up. Look at modules. Is dry deposition any good? Is chemistry good?

No.

No.

Runs from WOGA.

Air quality monitoring data trends. Compare to model predictions. Not bad for model. Check historically and it makes sense.

Ozone trends nonexistent. Are we doing any good? Maybe it would be worse without controls. Hard to relate trends and modeling. AIRSHED tends to corroborate what you would expect from analytical considerations.

No, except judgement that underlying principles are sound and it is the state-of-the-art model.

Whole business of UAM had to do with NO_x control. All perceptions based on answers by intuition. In 1979, NO_x control implied ozone up where it wasn't too bad, and down a bit where it was bad. But ups and downs not very significant. Background has always been NO_x yields ozone. Knew where to look to see if whoever was doing the modeling might bias inputs. Have to look at all factors in evaluating model. Can't be too impressed by nice slides or computer with differential equations. Check against smog chambers. Now gathering data to evaluate models for more than ozone and NO_x in time and space: intermediates, tracer releases, etc. Need to develop a strategy for reducing particulate. Want to use models, but need to satisfy ourselves that model not "kidding us". Many degrees of freedom; modeling can't be just curve fitting. May have good emissions and chemistry, but don't have assessment of vertical profile. Lack of input data is a problem; inputs are critical in terms of your assessment of reliability. If model is flat in response to inputs, even more worried.

A good model (hydrocarbons, NO_x, etc.) can be used qualitatively. Can probably never be used quantitatively, unless maybe very small grid, much better data, many meteorological regimes. Otherwise too many guesses.

No.

IMPACT OF UAM ON INTERVIEWEES' THINKING

1985 Berg work demonstrates validity of 1982 concerns. Even if model estimates are only qualitative, given the sensitivity work, District and EPA are taking a counterproductive tact.

For 1979 AQMP, marginally affected because two other simple models (rollback) showed similar results. By 1986, influence was substantial but in a qualitative way. Relied on sensitivity runs.

For Rule 1135 and later activities, influence was substantial because model was able to demonstrate a cause-effect relationship, as opposed to statistical correlations.

For 1979 AQMP, there was a policy problem: the model was very complicated, and you couldn't calculate allowable emissions with it. Had to rely on EKMA to comply with the law. For 1982, there were a large number of model runs. Had to come up with allowable emissions. General case: tended to confirm need to reduce hydrocarbons, maybe some effect due to NO_x.

Reinforced.

In absence of enormous quantities of real-time data, model is best tool.

Little impact for Rule 1135 because so many unanswered questions raised by technical people.

Other than historical data, model was the only factor.

Reinforced.

In 1982, little influence of model because control estimates were so far beyond what you could achieve. In Kern, model was absolutely determining element, and implied need for both hydrocarbon and NO_x controls. Planning to use model for Santa Barbara SIP.

Sophisticated models did not show NO_x impact. EKMA (based on lab experiments) showed NO_x impact. NO_x more expensive to control than hydrocarbons. Since federal standards for NO_2 violated, important to reduce NO_x independent of ozone. Big argument: does NO_2 have less impact on health than ozone? But Board focused on standards. EKMA said also good for reducing ozone.

Kern, dramatically.

Little impact, but didn't ignore. More recently, marginally; more confidence.

More limited since 1979 AQMP.

1982: moderate; set back those who thought NO_x control would help ozone. Will take enormous amount of precursor control.

INTERVIEWEES' ASSESSMENTS OF FACTORS AFFECTING INFLUENCE OF UAM

For gaining confidence, understand limitations of model. Must take variability into account. Numerical solutions have uncertainties. Main way to improve confidence: data base and build protocol that lays out how to deal with performance evaluation and what is acceptable to technical people and decision makers. If decision makers comfortable with range, OK. If uncomfortable with +/- 20 percent, not OK.

Need better data, but greater number of evaluations is most important. Have evolved in knowledge of model's strengths and weaknesses. Sensitivity analyses show important parameters.

Better data base. Requires initial and boundary data. Can overcome lack of initial but need boundary. Can't get boundary aloft; can partly overcome by sensitivity analyses. Need application to different years with different emission levels and patterns.

Show more sensitivity results. Need better data but if model is sensitive to +/- 10 percent wind uncertainties and we can only measure to +/- 20 percent, there's a real problem. Need to know limits of data; either improve the data or don't model.

More influential if lesser sensitivity to input uncertainties; model eats up data.

Conversely, if sensitivity runs went to pot, would have shot confidence.

Might be better if not so sensitive to initial and boundary conditions, but may not be possible.

Ever since results first done, no new findings. Results already there. Rerun by WOGA didn't really cause them to reconsider.

Focus on ozone: impact of elevated versus ground-level emissions.

Plus knowing inadequacies.

Limitations: quality of input data, improvement to individual model subcomponents (weak links). Data upgrade. Greater independence of people who performed calculations.

Reduce underprediction: operationally difficult, requiring a lot of data. Does not have latest chemical mechanism; no assessment if performance would change. Initial conditions potentially a concern but impression was they weren't as important. Should take a closer look at hydrocarbon-to-NO_x ratio.

Now officially sanctioned; in guidelines as recommended model. Better data would yield more confidence in model. Greater number of evaluations: tradeoff cost versus availability. Now on PC's.

Try to see if input data would need to be sharpened up.

Model where it ought to be.

Multi-day: overdependency on initial conditions. Previous-day contribution more of an artifact of model than a real phenomenon. Not too bad for applying model, since don't do day-to-day. Not much concern for inputs, but more for chemistry, speciation. Doing more evaluations now. Need more episodes. If results show consistency, will have more confidence.

Greater degree of accuracy and reliability not only for ozone but also for intermediates. Hydrocarbon-to-NO_x data.

Excellent job of ozone. Hydrocarbon-to-NO_x ratio: may be a problem with chemistry, should be comparable with measurements, assuming they are correct.

Better emissions data.

Better data for inputs (precursors data). Santa Barbara - Ventura: three to five meteorological regimes; rely more on measured data and less on emissions.

INFLUENCE OF UAM ON INTERVIEWEES' ASSESSMENTS OF SPECIFIC ISSUES

Now trying to assess utility for single-source determination. Trajectory models from point sources can overestimate; puff disperses in real world more than in model.

All applications reinforced opinion that grid-based model useless for single-source determination.

Later PARIS. Don't see major differences between PARIS and AIRSHED. Some differences in near field. Either adequate. PARIS more expensive.

Biggest problem: policy makers think you can do it but you can't because no single source is large enough.

Not an issue: largest NO_x source only about one percent of emissions.

Need data base with finer meteorological and gridded emission inventory.

If you don't have data, don't model.

Model raised issues and forced thinking.

Completely convinced that the more clearly the data are designed for the model, the better off you are.

Had previously thought about effects of NO_x controls because of uncertainty. Some locations with NO_2 problem. Very important to analyze. Highly controversial because of concept that decreasing NO_x might increase ozone. Model gave insights into what might be happening. Useful for both NO_2 and ozone impacts. Designed data base would have been and continues to be desirable; need data that are current for emissions and meteorology. Validate more recently to increase confidence. Don't usually look at single sources, only single categories at various locations.

Feel strongly that grid-based model is not appropriate for single-source determination. Uncertainty in model is too high; lost in noise.

Designed data base nice if you can afford it. Wish model were good for single sources but know SAI couldn't make it so.

Resolution of source strengths on grid was limiting; needed better data than you had. Well worth having for whole regulatory process.

Model reinforced things already understood.

Don't have that degree of resolution.

Model extremely helpful for effects of NO_x on ozone.

So many sources, so widespread; too difficult to detect single sources.

Model does show NO_x control affects ozone. Gross mistake to go someplace without intensive aerometric data base. No on single source determination, very strongly.

Ideally run for multi-day period. In principal, recommend multi-day. Kern SIP: single-day because had problem with hydrocarbon-to-NO_x ratio. Data bases obviously very important. Most routine data bases unsuitable. Stations often near local sources. No upper air data. If you have upper air, you will get better performance. Have done a number of runs for single source. Don't see much impact from single source, as you expect. If you do see an impact, in 1-5 ppb range. ARB management does not believe that photochemical models can be used for single source.

Don't see single source; would never use for single source.

No - I remain skeptical on single source.

Don't yet know answer to question of single source impact. Will learn a lot from source categories.

Data base critical. May be 3-4 applications, requiring better and better data bases. Not useful for assessing single sources - lack of precision and accuracy.

Impacted thinking on NO_x controls. Basic strategies changed to NO_x. Model verifies what you could have guessed from hydrocarbon-to-NO_x ratio.

INTERVIEWEES' ASSESSMENTS OF GENERAL INFLUENCE OF UAM

Others weren't ready to accept and believe that NO_x controls detrimental. EPA Region IX assertive that controls needed. Spatio-temporal aspects ignored. Evolution: other ways of looking at ozone unacceptable (EKMA worthless and trajectory can't work).

Kern, the strongest influence, but didn't focus anyone.

Internally, quite substantial. Results surprising: reduce NO_x and ozone levels in near field increase. Actual long-term measurements have verified trends. Results obtained led to regulatory inertia; results contradictory to preconceived position. Made regulatory agencies think about results but couldn't accept them. Local District using AIRSHED for 1987 plan.

Models too important to leave to public agencies. See District and ARB relying on models - want to make sure not misapplied. Industry now spending money to support District efforts.

Substantial effect but not realized immediately. Regulatory agencies defended their own plan because done prior to model. By 1985, regulatory agencies definitely influenced in thinking. Confidence seems to go up with each iteration. Large skepticism on part of regulatory community because used by industry. Probably as it should be. Now used by regulatory community.

Industry affected by UAM but not ARB, SCAQMD (could not accept concept of NO_x).

For those who agreed with results, made them more entrenched. For those who disagreed, made them distrust models. Now, resurgence in statistical approaches. According to regulatory community, statistical approaches describe reality. Model didn't support their position, so now use statistical models. Can draw correlations, but cause-effect may not be there. Difficult in complex area like Los Angeles to mistake cause-effect.

Model raised more questions than it answered. Same with 1979 and 1982 AQMP's. Decision makers had a hard time coming to grips with model. How do you translate data into a usable decision-making format? Easy to use EKMA or rollback. Model doesn't say in black and white what you need to do.

Berg hearings were two to three members of District Board interacting with about ten scientific people. Initially, question was should NO_x be reduced to reduce dry acidic deposition. Became an argument over effect of NO_x on ozone. Only people whose opinions counted were District Board members. Impression was that Board members came in with preconceived opinions. Looking for scientific support, not information (not applicable to scientists).

In 1982, everyone gasped when they looked at the bottom-line numbers. Model was absolutely critical for Kern.

In 1979 and 1982, little influence on Board, little heavier influence on District.

Complete puzzlement on part of decision makers. Nonscientific community incredulous at how little we know about modeling. Don't want to squeeze more out of model than we should.

Little influence on Coalition for Clean Air; substantial influence on District.

Influence waxed and waned, depending on Board members and degree of external pressure.

INTERVIEWEES' ASSESSMENTS OF FACTORS OF MORE INFLUENCE THAN THE UAM

First, as you approach 100 percent reductions, modeling process questionable. Second, social issues implied by large emission reductions. Third, debate between hydrocarbon and NO_x control. Fourth, cost.

Politics.

Rule 1135, NO_x control for NO_2 compliance. Kern county, battle over modeling.

Notion that any emission reduction is good. Regulators use dollars per ton of pollutant removed. Should use dollars per air-quality improvement.

First, costs of controls and modeling runs. Second, EPA requirements: meeting standards and what kind of modeling.

Board got conflicting testimony from staff and industry; came down with staff. People who are going to be affected have to be involved, have to have opportunity to provide you with information along the way.

First, people's predetermined positions were more important. Second, existing bureaucracy - District and ARB clash. Third, views on air pollution. Legally, a health hazard, but public sees it as a nuisance; attitude plays into view of model.

Value judgments about what you do in face of uncertainty. Do you opt to go slow on perhaps costly regulations? Do you push ahead because you want a margin of safety?

Politics: there is more to regulation than science.

Need to get decision makers to buy into process without prejudice of knowing the results.

Gross factor: any emissions, no matter what pollutant or where from, endangered public health; therefore, any reduction was good, whether model said it was good or not. True environmentalists: pound reduced anywhere worth doing. Would challenge any model that implied selective reduction. Industry feels the other way (too much so). Bias: easier to cut emissions from large industry then from small sources that affect the public directly (e.g., space heaters, leaf blowers, etc.). Requiring refinery to install controls is applauded by the public. But if model said the opposite, it wasn't supported by the political people. Model told you controlling heavy industry wasn't sufficient. But public rejects direct costs. Legislature gives District complete authority over industry but challenges any rules affecting constituents.

Integration of information sources; hard to separate. If model is consistent with the mainstream of experience, people comfortable with it and willing to extrapolate.

Personal values of decision makers (e.g., one Board member may be a fire-breathing environmentalist and another would prefer to see no money laid out). Models don't make decisions, but they do raise the level of discussions. You get a more focused discussion. Those favoring controls focus on downwind. Those opposing controls focus on near-field ozone.

Both EKMA and AIRSHED influential.

EKMA and federal NO_2 standards (exceeded based on real measurements); get NO_2 exceedances down.

Pre-existing biases and innate convictions. Propaganda - people who have fundamental beliefs are not subject to analytical discussion. Model is helpful in getting people to think about ideas.

People try to get sense of historical trend in air quality with respect to emissions. Always preconceived as solid evidence. Unfortunately, you can't separate effects of various pollutants since you only have the actual scenario. Don't see a contradiction between modeling and real data. Decision makers have to be told that what happened in past is the combined result of a lot of activities. Danger of using two different sources of information when they really serve two different functions. Historical data shows the effect of hydrocarbon and NO_x decreases but can't show other scenarios. Therefore, you can't tell for the future what would result if you just controlled hydrocarbons or just controlled NO_x. Probably makes sense to predict the future from trends if control scenario similar to past.

Political influences, considerations of cost, and lobbying.

Hydrocarbon-to-NO_x ratio (high, greater than 15-20:1). Knowledge of chemistry implies that NO_x decreases are beneficial. Model is not the only game in town. Combination of modeling, analyzing ambient data, and common sense.

Decision makers worry not just about technical issues. Final agreed-upon SIP reductions had more to do with politics and compromise. EKMA and linear rollback were very similar in Los Angeles: very comforting to policy makers.

First, Jim Pitts' and Art Weiner's questioning of correlation of model to reality; second, ARB's statistical information; SCAQMD staff's reservations.

Group consensus of Board changes on whether to pay attention to model or go with political winds.

In 1979, issue was amount of control that could realistically occur. Model wasn't important. All tools said 80 to 90 percent control. Could control measures get that much control? In 1982, predicted ways of not attaining standards; model not important. Kern county, model was critical.

Hydrocarbon-to-NO_x ratio (a piece of the model - may not need full-blown model).

ASSESSMENTS OF USES OF MODELS AND TRENDS IN APPROPRIATE USES

Model became a diagnostic tool in the 1980's: take a group of measures and measure impact.

By 1980, appropriate for assessing outcomes, but you have to look at other information. Have you seen what model predicted earlier? Do models make sense? Other information, diagnostics.

In 1979, thought that if flaws were fixed, models could be a useful tool. Now think they are useful only for gross sensitivity analyses. Frightening to see them used for individual projects. The more we see applications to specific projects, the more our worst fears are true: some technician making some initial assumption that drives conclusions.

By 1980, good enough as assessment tool in relative sense: how different scenarios might rank, which ones get you to standards in the most cost-effective manner. Regulatory agencies made a change since 1980.

Understanding developed between 1976 and 1980; became a directional tool. Since 1980, more appropriate uses for model, more understanding on part of staff and Board.

In 1975, EPA and ARB couldn't deal with complexities. Clean Air Act said you shall have a plan to deal with the problem as simply as possible. Now not dealing with simplification. What's changed is not the model but recognition that those models are the only tools to deal with complexity of the problem. Can't just reduce ROG to meet ozone standard. Models haven't evolved that much; models may be the only constant - everything else has changed. Toxic versus nontoxic ROG's. Realization that models can deal with complexity. Big change is that now they're cheap to run. Collecting emissions data for input to model; more efficient at getting emissions data into model. Integration of effort.

Not much change between 1975 and 1980. 1980 to 1982, looking at sensitivity analyses. Now appropriate for assessing outcomes.

As of 1981 - 1982, not good enough for assessment.

Models require more information than is required for real-life decisions. May help identify possible unexpected results. Least successful for assessing outcomes of control measures. No perceived major change between 1975 and 1980. Since 1980, enough people to exercise model so that it became important to look at models and use information. Could or should they get to the point of exclusive reliance? Dubious; need to have other factors - moral, ethical, and political regardless of technical information.

Model tells you where you need more data.

Further refinement between 1975 and 1980; more useful as a tool. It is a tool, but it's not the only tool.

Better science between 1975 and 1980.

Between 1975 and 1980, better understanding of what models could and couldn't do (period of competition between SAI and Cal Tech). Since 1980, understanding of models elevated from technical and policy staff to Boards.

In 1975, barely acceptable. In 1980, accepted as going into AQMP's. Both changes in model and data bases and people's confidence in model.

Technology of performing calculations hasn't changed much. Distressingly little difference between 1975 and 1986. Resource limited. Differences have to do with accumulation of input data and results. Issue is people, not technology. Should have outlined crash research program in 1973. Distressingly little use. Reason is political, not technological. Interesting to note that a large fraction of air-pollution decision makers do not want their lives complicated by models. Ability to compromise is diminished by the facts. If you know what you're up against, ways to make models effective: menu of alternatives. Define alternatives in terms of costs and consequences. Cannot give a simple answer. Doesn't take prerogative to make decision from decision maker.

State of art improved from 1982 to 1986. Better chemistry (nighttime NO_x removal). Better data bases. Pay more attention to performance evaluations. Change in regulated community. Now willing to spend money to collect data. Regulating agencies more willing to entertain model results and put up money to collect data. In past, just ran EKMA, even though everyone knew it was wrong, and looked the other way.

In 1976, would encourage progress. Confident that if model could come up to expectations, would be useful. Input data always the same. Treatment of data with respect to species improved. Kinetic data needed. Continuing effort to upgrade winds, point sources, etc. Between 1979 and 1982, probably improved, but not enough to swing confidence. Area of improvement: cut down on time needed for running. Make it easier to use model.

In 1977, no need for models: no data, no models. Few large emitters. Simple rollback could do the job. Between 1975 and 1980, assess trends as a function of strategies. Now trying to assess outcomes of measures of policy decisions. Requires higher confidence. Now dealing with implications of population growth or transportation. Implications of zero growth very important.

Since 1980, changed from a research tool to a regulatory tool.

Since 1980, use qualitatively to complement other analytical tools.

In 1975, ARB at square zero. Had not used any photochemical models, not even Gaussian. Only line-source models. Beginning 1976, ARB funded Cal Tech 3D model. Didn't get there with Cal Tech model. Also, interest in LIRAC in Bay area. Kern county classic example of identifying possible outcomes requiring further investigation.

In 1975, at such an early stage. Thought maybe models could be used to assess changes in land-use planning (spatial and temporal shifts). Between 1975 and 1980, models became marginally more believable. Still a long way to go. Since 1980, would still like to use models to assess land-use measures. Did some testing using UAM for land-use changes. Model predicted no change, but 1984 Olympics showed air-quality change.

Originally thought you could use models conclusively to show non-utility of NO_x controls.

In 1975, no substantial use for influencing policy. Began to see use between 1975 and 1980; not aware of improvements. Important changes between 1980 and 1985; have gone from believing models magic. Now coming out of backlash of thinking models junk. More balanced and realistic view. If properly done and utilized, can be useful. Improvements in model have helped.

ADDITIONAL COMMENTS

Models extra complicated and expensive in development and use. Should be continued. Reason: ability to model demonstrates an understanding and forces scientists to understand processes. Without demonstration of understanding, political and public support cannot be developed. In 1975, could just say air is bad and have to clean it up. Now, must have a precise and manageable tool to determine where to go and how to get there. If you can't say anything after 30 years, have to give up.

Focus on trying to destroy model's results is counterproductive. Before you do anything with model, all groups must establish first, what they're looking for, second, acceptable uncertainties, and third, a protocol that lays out how you do performance evaluations.

Planning board set up to advise District on use of models: representatives from WOGA, SCE, and District.

Not sure that Boards have knowledge of complexities. Greater access to computers, better data (better emissions data, location of sources), nighttime chemistry slowly but surely improved the model, and have made it a useful planning aid. Just acquired computer with model installed. Can make more runs and more sensitivity analyses. Limited before because couldn't afford to make more than three runs. Limiting factor: model doesn't cover the entire basin; sensitivity to boundary conditions and paucity of data.

Everyone wants models, but no one is willing to put money in to make model work great. No one wants to go back to best control technology. Use model to say how far to go.

Want model to be comprehensive. Most useful for indicating directions of changes. Pleased that acid deposition modeling is done under government auspices.

SCCCAMP: effort to collect data needed to drive model, between 1982 and 1985. Increased confidence between 1982 and 1985. Most evolution in thinking occurred in short period in 1982. SCCCAMP strengthened. Cooperative groups (technical and policy) from regulatory, regulated, academic, and environmental sides (last may not have technical expertise). Takes time and effort to overcome distrust and show that everyone is working toward a common goal. Then get technical details together. Necessary because expensive - takes time to get money into budgets. In long run, in society's interest; adversarial roles not constructive. Tried a joint venture in 1982, but too much rancor and skepticism. Industry did it themselves, but viewed with so much skepticism by regulators that they discounted it. Then went to counties, state, and EPA. Proposed sitting down and figuring out best way to answer questions. At least, how to collect data so that you agree the data are satisfactory. Ultimately want to agree on modeling protocol, but at least agree on the data. Took about one year. How you do modeling is still open to question, but now agree that modeling is to be done. With respect to acid deposition, need to agree on data, then agree on need to use a model. Get concerns on the table and try to hammer them out. With SCCCAMP, found much posturing and much argument tied up in single concerns. Thought other entity out to gore your ox when really other guy was trying to protect two or three little things that may not be of concern to you. Found out that 90 percent of stuff that you thought you had huge differences over really weren't. Policy agreement is needed before technical guys start. Then let the technical people go ahead (they usually have few differences). Can't happen in three months. Need time and effort to build trust. SCCCAMP hardest thing ever done. Will be hard to use model in as totally cooperative a committee as was done with the data collection because regulatory agencies concerned that industry is using model to its advantage. Properly concerned that governmental entities should be doing modeling so that industry is not dictating policy. Appropriate for industry to be aware of and have a chance to make additional runs (maybe with different parameters). Doesn't compromise government and gives industry a chance to have concerns addressed. Built an agreement over data. Not cheap and not quick, but can be done. Convinced that this modeling is a good way to do things and is the best tool now.

Familiar with models from District and industry perspective. People have preconceived notions. If model supports it, model is valid. If model doesn't, reaction is "model can get you any answer you want". In actuality, if results disagree with preconceptions, you say study is no good. End of story is "I don't believe model". In non-scientific community, people often say I just disagree - my gut says so. Decision-making body is non-technical. Can exercise model all you want, and remains unconvincing. Decision maker uses results of models partly because they support his case. Need to present other evidence because decision-making community can't rely only on model. If "druthers", have agency buy into an analytical mechanism that would make results binding on both sides. SAI model is only game in town; if criticism of assumptions gotten, it would be helpful. Relatively costly control measures save in pphm range. Industry would argue why spend e.g. 500 million to change ozone from 0.36 to 0.357 ppm? People opposed to model may fear that model will reveal this outcome; not going into model with open mind. Public should know dollars per gain. Cornerstone of new study framework, which District approved three weeks ago, is to use dollars per gain (ppm) improvement, not dollars per ton removed. You have to face that Los Angeles won't make 1987 ozone standard. Do you shut down the city or do something reasonable? Need to know what you're buying - improvement in air quality (protection of public health and welfare) - not in terms of tons pollutant removed. This helps focus debate on defining what is reasonable. One possible approach might be probabilistic, using decision tree. No decision is 100 percent right or wrong. Allows people to drop guard somewhat. Focus discussion on assumptions of model in probabilistic way. Some assumptions may not be as sensitive to probabilistic variance. Those that are warrant investigation. Do sensitivity analyses in probabilistic manner.

View models as a planning tool. Not definitive. Can give valuable information directionally. Can give relative changes of 20 to 30 percent. Couldn't use to make decision on project approval. One issue is that model only predicts for data that you give it. May not have a representative day (ozone-wise) for meteorology. Performance needs to improve for predicting peak ozone levels. Models may perform well on average but not on peaks. Tweak model to get performance on peaks, then average performance goes to hell. Uncertainties too high for single source; may never be resolved.

Money for modeling should come from government, not industry, to avoid perception of tainting. University problem: can't develop a model on time and budget.

Absolutely critical issue is perceived independence of people doing the work from both industry and government pressure. Need system that insulates people doing the work from repercussions. If one side captures all the modeling expertise, that side can use technical expertise to club the other side. Results are on table for all to use; both sides pull out what they want. Gets ugly if one side has bought up all the technical talent. Models are complicated; few can deal with them. Possible to monopolize talent. Focus not on technology but on process of using results. Technological problems more easily resolved than institutional problems. To add confidence to the SAI model, things you do would be different than what you do to NAPAP model on first time out. For NAPAP, need to get rid of initial technical problems, module problems. Rude awakening to need for good emissions inventory. Big problem: job of emissions inventory is unglamorous. In Los Angeles, modeling is one part model to two parts straightening out emissions.

PARIS model: if very large SO_2 sources, might be able to use a gridded model. For ozone, too many very small sources.

Little point to run a complex model if you don't have a feeling that out of it will come a better quantification than with a simpler model. Not sure if crossover has yet occurred. Confident that it will with time but it hadn't during 1979 - 82 period. As regulator, the question is: am I spending money to move toward a desirable goal? Want to feel that you're not moving in the wrong direction when spending money. Had to take stringent opinions if dealing with carcinogens or mutagens (e.g., vinyl chloride). Negativity concerning past models may not be fair relative to current models.

Would be nice to have a "sales brochure" for model at a reasonable technical and policy level, e.g., here is model, what it can do, how you can use it. Would use more if available.

AIRSHED: a lot of work on how to make model predict best, beefing up chemistry, etc. Needs documentation on input structure and how to best use model: data needs, plan for updates, what are future plans, modification protocol, how to use model, how many episodes you need to cover, sensitivity, performance. Many people don't know how much resources required to run model. Input structure is more complex than model itself.

For future, no other way besides airshed modeling. Procedures will be as or more comprehensive than for Kern county. Everyone learned a tremendous amount about model in Kern county application. Next major effort will be Santa Barbara, where you have SCCCAMP data base. Will set higher standards since SCCCAMP very comprehensive.

Need a lot of checkpoints along the way. Any modeler can reproduce ozone or acid deposition. But you must reproduce all the stuff in between. Long way to go for acid deposition modeling. A million square miles compared to 2000 square miles in Los Angeles. Atmospheric chemists still argue about what's going on in the atmosphere. Do modeling anyway. Don't buy argument that you have to sit around to achieve perfect understanding before you do anything. Cannot keep going back and saying that you need a better dispersion model, etc., and keep putting off decisions. Problems with UAM will be confounded one to two orders of magnitude in acid deposition modeling. Emissions, geographical size, ammonium, calcium, transport, chemistry in clouds. RADM is rain event basis but effects are long-term average. Response to criticism is to run RADM for 365 days. Probably won't work. Also RADM OK for wet deposition but what about dry and fog? RADM has left out field study part. Have struggled at ARB to see what to do about acid deposition models. Cost-effectiveness of controls should be main issue, not developing absolute source-receptor relationships. Shouldn't buy time (avoiding the real issue). Cannot not make decisions. Do best work and still sponsor. Would be here forever if tried to do modeling perfectly. More difficult questions are technical and cost-effectiveness.

Developed skepticism. Experts raised questions. Deep effect this past year. In relationship to NO_x hearings, heard ARB presentation on statistical correlations between NO_x emissions and ozone downwind. Model probably delayed air quality control. Model did not correspond with later statistical information. Maybe with adjustments it will be consistent. Skeptical of model so subject to small errors. Concern about fine-tuning when so far away from attainment. May be appropriate when close and you need to hone down.

Need to use models in a way in which regulators/regulated work together to evaluate best strategies. To meet ozone standard anytime soon is completely unrealistic goal. Models imply need to reduce vehicle miles traveled by 50 percent and remove all stationary sources. Contrary to Clean Air Act goals. Doesn't mean should do nothing. As policy makers, ask what interim goals with what effort. Ask what we want. Think that SCE wants plan resulting in greatest lessening of health risk at lowest cost. Will force thinking of innovative strategies (time and place controls). Need to all work together to upgrade models and data inputs. Change since 1975: now beginning to have confidence that you can do it. Supplement with case studies (e.g., during Olympics, about 14 percent ozone reduction by traffic controls). Have to see how politicians will answer.

Dilemma for EPA. Clean Air Act requires demonstration of attainment. Have to do so with models, but models aren't good enough. Model only good enough for directionally-correct decisions. Demand placed on EPA to demonstrate attainment; CARB quit asking models to do that. Clean Air Act requirement will survive. Continuing legislative demand for models to do something they can't do. Congress should realize this and redesign Clean Air Act. Models will get better but will never be totally precise. Input data more important than model itself. Congress won't recognize inadequacy of models and restructure Clean Air Act.